Advances in Applied Mathematics

FOURTH EDITION

Advanced Engineering Mathematics with MATLAB®

Dean G. Duffy

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

FOURTH EDITION

Advanced Engineering Mathematics with MATLAB

Dean G. Duffy

CRC Press

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper
Version Date: 20161121

International Standard Book Number-13: 978-1-4987-3964-1 (Hardback)

Library of Congress Cataloging-in-Publication Data

Names: Katzman, Steve (IT auditor)
Title: Operational assessment of IT / Steve Katzman.
Description: Boca Raton, FL : CRC Press, 2016. | Series: Internal audit and IT audit ; 4 | Includes bibliographical references and index.
Identifiers: LCCN 2015037136 | ISBN 9781498737685
Subjects: LCSH: Information technology--Management. | Information technology--Auditing. | Auditing, Internal.
Classification: LCC HD30.2 .K3857 2016 | DDC 004.068--dc23
LC record available at http://lccn.loc.gov/2015037136

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Chapter 12: The Laplace Transform 559

Chapter 13: The Z-Transform 667

Introduction

Today's STEM (science, technology, engineering, and mathematics) student must master vast quantities of applied mathematics. This is why I wrote *Advanced Engineering Mathematics with MATLAB*. Three assumptions underlie its structure: (1) All students need a firm grasp of the traditional disciplines of ordinary and partial differential equations, vector calculus, and linear algebra. (2) The digital revolution will continue. Thus the modern student must have a strong foundation in transform methods because they provide the mathematical basis for electrical and communication studies. (3) The biological revolution will become more mathematical and require an understanding of stochastic (random) processes. Already, stochastic processes play an important role in finance, the physical sciences, and engineering. These techniques will enjoy an explosive growth in the biological sciences. For these reasons, an alternative title for this book could be *Advanced Engineering Mathematics for the Twenty-First Century*.

This is my fourth attempt at realizing these goals. It continues the tradition of including technology into the conventional topics of engineering mathematics. Of course, I took this opportunity to correct misprints and include new examples, problems, and projects. I now use the small rectangle □ to separate the end of an example or theorem from the continuing text. The two major changes are a section on conformal mapping (Section 10.11) and a new chapter on stochastic calculus.

A major change is the reorganization of the order of the chapters. In line with my goals I have subdivided the material into three groups: classic engineering mathematics, transform methods, and stochastic processes. In its broadest form, there are two general tracks:

Differential Equations Course: Most courses on differential equations cover three general topics: fundamental techniques and concepts, Laplace transforms, and separation of variable solutions to partial differential equations.

The course begins with first- and higher-order ordinary differential equations, Chapters 1 and 2, respectively. After some introductory remarks, Chapter 1 devotes itself to presenting general methods for solving first-order ordinary differential equations. These methods include separation of variables, employing the properties of homogeneous, linear, and exact differential equations, and finding and using integrating factors.

The reason most students study ordinary differential equations is for their use in elementary physics, chemistry, and engineering courses. Because these differential equations contain constant coefficients, we focus on how to solve them in Chapter 2, along with a detailed analysis of the simple, damped, and forced harmonic oscillator. Furthermore, we include the commonly employed techniques of undetermined coefficients and variation of parameters for finding particular solutions. Finally, the special equation of Euler and Cauchy is included because of its use in solving partial differential equations in spherical coordinates.

Some courses include techniques for solving systems of linear differential equations. A chapter on linear algebra (Chapter 3) is included if that is a course objective.

After these introductory chapters, the course would next turn to Laplace transforms. Laplace transforms are useful in solving nonhomogeneous differential equations where the initial conditions have been specified and the forcing function "turns on and off." The general properties are explored in Section 12.1 to Section 12.7; the actual solution technique is presented in Section 12.8.

Most differential equations courses conclude with a taste of partial differential equations via the method of separation of variables. This topic usually begins with a quick introduction to Fourier series, Sections 5.1 to 5.4, followed by separation of variables as it applies to the heat (Sections 8.1–8.3), wave (Sections 7.1–7.3), or Laplace's equations (Sections 9.1–9.3). The exact equation that is studied depends upon the future needs of the students.

Engineering Mathematics Course: This book can be used in a wide variety of engineering mathematics classes. In all cases the student should have seen most of the material in Chapters 1 and 2. There are at least four possible combinations:

• *Option A*: The course is a continuation of a calculus reform sequence where elementary differential equations have been taught. This course begins with Laplace transforms and separation of variables techniques for the heat, wave, and/or Laplace's equations, as outlined above. The course then concludes with either vector calculus or linear algebra. Vector calculus is presented in Chapter 4 and focuses on the gradient operator as it applies to line integrals, surface integrals, the divergence theorem, and Stokes' theorem. Chapter 3 presents linear algebra as a method for solving systems of linear equations and includes such topics as matrices, determinants, Cramer's rule, and the solution of systems of ordinary differential equations via the classic eigenvalue problem.

• *Option B*: This is the traditional situation where the student has already studied differential equations in another course before he takes engineering mathematics. Here separation of variables is retaught from the general viewpoint of eigenfunction expansions. Sections 9.1–9.3 explain how any piece-wise continuous function can be reexpressed in an eigenfunction expansion using eigenfunctions from the classic Sturm-Liouville problem. Furthermore, we include two sections that focus on Bessel functions (Section 6.5) and Legendre polynomials (Section 6.4). These eigenfunctions appear in the solution of partial differential equations in cylindrical and spherical coordinates, respectively.

The course then covers linear algebra and vector calculus as given in Option A.

• *Option C*: I originally wrote this book for an engineering mathematics course given to sophomore and junior communication, systems, and electrical engineering majors at the U.S. Naval Academy. In this case, you would teach all of Chapter 10 with the possible exception of Section 10.10 on Cauchy principal-value integrals. This material was added to prepare the student for Hilbert transforms, Chapter 14.

Because most students come to this course with a good knowledge of differential equations, we begin with Fourier series, Chapter 5, and proceed through Chapter 14. Chapter 11 generalizes the Fourier series to aperiodic functions and introduces the Fourier transform. This leads naturally to Laplace transforms, Chapter 12. Throughout these chapters, I make use of complex variables in the treatment and inversion of the transforms.

With the rise of digital technology and its associated difference equations, a version of the Laplace transform, the z-transform, was developed. Chapter 13 introduces the z-transform by first giving its definition and then developing some of its general properties. We also illustrate how to compute the inverse by long division, partial fractions, and contour integration. Finally, we use z-transforms to solve difference equations, especially with respect to the stability of the system.

Finally, there is a chapter on the Hilbert transform. With the explosion of interest in communications, today's engineer must have a command of this transform. The Hilbert transform is introduced in Section 14.1 and its properties are explored in Section 14.2. Two important applications of Hilbert transforms are introduced in Sections 14.3 and 14.4, namely the concept of analytic signals and the Kramers-Kronig relationship.

•*Option D*: Many engineering majors now require a course in probability and statistics because of the increasing use of probabilistic concepts in engineering analysis. To incorporate this development into an engineering mathematics course we adopt a curriculum that begins with Fourier transforms (minus inversion by complex variables) given in Chapter 11. The remaining portion involves the fundamental concepts of probability presented in Chapter 16 and random processes in Chapter 17. Chapter 16 introduces the student to the concepts of probability distributions, mean, and variance because these topics appear so frequently in random processes. Chapter 17 explores common random processes such as Poisson processes and birth and death. Of course, this course assumes a prior knowledge of ordinary differential equations and Fourier series.

A unique aspect of this book appears in Chapter 18, which is devoted to stochastic calculus. We start by exploring deterministic differential equations with a stochastic forcing. Next, the important stochastic process of Brownian motion is developed in depth. Using this Brownian motion, we introduce the concept of (Itô) stochastic integration, Itô's lemma, and stochastic differential equations. The chapter concludes with various numerical methods to integrate stochastic differential equations.

In addition to the revisions of the text and topics covered in this new addition, MATLAB is still employed to reinforce the concepts that are taught. Of course, this book still continues my principle of including a wealth of examples from the scientific and engineering literature. The answers to the odd problems are given in the back of the book, while worked solutions to all of the problems are available from the publisher. Most of the MATLAB scripts may be found at http://www.crcpress.com/product/isbn/9781439816240.

Definitions

Function	Definition
$\delta(t-a)$	$= \begin{cases} \infty, & t = a, \\ 0, & t \neq a, \end{cases} \qquad \int_{-\infty}^{\infty} \delta(t-a)\,dt = 1$
$\operatorname{erf}(x)$	$= \dfrac{2}{\sqrt{\pi}} \int_0^x e^{-y^2}\,dy$
$\Gamma(x)$	gamma function
$H(t-a)$	$= \begin{cases} 1, & t > a, \\ 0, & t < a. \end{cases}$
$H_n^{(1)}(x), H_n^{(2)}(x)$	Hankel functions of first and second kind and of order n
$\Im(z)$	imaginary part of the complex variable z
$I_n(x)$	modified Bessel function of the first kind and order n
$J_n(x)$	Bessel function of the first kind and order n
$K_n(x)$	modified Bessel function of the second kind and order n
$P_n(x)$	Legendre polynomial of order n
$\Re(z)$	real part of the complex variable z
$\operatorname{sgn}(t-a)$	$= \begin{cases} -1, & t < a, \\ 1, & t > a. \end{cases}$
$Y_n(x)$	Bessel function of the second kind and order n

Chapter 1
First-Order Ordinary Differential Equations

A *differential equation* is any equation that contains the derivatives or differentials of one or more dependent variables with respect to one or more independent variables. Because many of the known physical laws are expressed as differential equations, a sound knowledge of how to solve them is essential. In the next two chapters we present the fundamental methods for solving *ordinary differential equations* - a differential equation that contains only ordinary derivatives of one or more dependent variables. Later, in Sections 11.6 and 12.8, we show how transform methods can be used to solve ordinary differential equations, while systems of linear ordinary differential equations are treated in Section 3.6. Solutions for *partial differential equations*—a differential equation involving partial derivatives of one or more dependent variables of two or more *independent* variables—are given in Chapters 7, 8, and 9.

1.1 CLASSIFICATION OF DIFFERENTIAL EQUATIONS

Differential equations are classified three ways: by *type*, *order*, and *linearity*. There are two *types*: *ordinary* and *partial differential equations*, which have already been defined. Examples of ordinary differential equations include

$$\frac{dy}{dx} - 2y = x, \tag{1.1.1}$$

$$(x - y)\,dx + 4y\,dy = 0, \tag{1.1.2}$$

$$\frac{du}{dx} + \frac{dv}{dx} = 1 + 5x, \tag{1.1.3}$$

and

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y = \sin(x). \tag{1.1.4}$$

On the other hand, examples of partial differential equations include

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0, \tag{1.1.5}$$

$$y\frac{\partial u}{\partial x} + x\frac{\partial u}{\partial y} = 2u, \tag{1.1.6}$$

and

$$\frac{\partial^2 u}{\partial t^2} + 2\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}. \tag{1.1.7}$$

In the examples that we have just given, we have explicitly written out the differentiation operation. However, from calculus we know that dy/dx can also be written y'. Similarly the partial differentiation operator $\partial^4 u/\partial x^2 \partial y^2$ is sometimes written u_{xxyy}. We will also use this notation from time to time.

The *order* of a differential equation is given by the highest-order derivative. For example,

$$\frac{d^3y}{dx^3} + 3\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^2 - y = \sin(x) \tag{1.1.8}$$

is a third-order ordinary differential equation. Because we can rewrite

$$(x+y)\,dy - x\,dx = 0 \tag{1.1.9}$$

as

$$(x+y)\frac{dy}{dx} = x \tag{1.1.10}$$

by dividing Equation 1.1.9 by dx, we have a first-order ordinary differential equation here. Finally

$$\frac{\partial^4 u}{\partial x^2 \partial y^2} = \frac{\partial^2 u}{\partial t^2} \tag{1.1.11}$$

is an example of a fourth-order partial differential equation. In general, we can write an nth-order, ordinary differential equation as

$$f\left(x, y, \frac{dy}{dx}, \cdots, \frac{d^n y}{dx^n}\right) = 0. \tag{1.1.12}$$

The final classification is according to whether the differential equation is *linear* or *nonlinear*. A differential equation is *linear* if it can be written in the form:

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = f(x). \tag{1.1.13}$$

Note that the linear differential equation, Equation 1.1.13, has two properties: (1) The dependent variable y and *all* of its derivatives are of first degree (the power of each term involving y is 1). (2) Each coefficient depends only on the independent variable x. Examples of linear first-, second-, and third-order ordinary differential equations are

$$(x+1)\,dy - y\,dx = 0, \tag{1.1.14}$$

$$y'' + 3y' + 2y = e^x, \tag{1.1.15}$$

and

$$x\frac{d^3y}{dx^3} - (x^2+1)\frac{dy}{dx} + y = \sin(x), \tag{1.1.16}$$

respectively. If the differential equation is not linear, then it is *nonlinear*. Examples of nonlinear first-, second-, and third-order ordinary differential equations are

$$\frac{dy}{dx} + xy + y^2 = x, \tag{1.1.17}$$

$$\frac{d^2y}{dx^2} - \left(\frac{dy}{dx}\right)^5 + 2xy = \sin(x), \tag{1.1.18}$$

and

$$yy''' + 2y = e^x, \tag{1.1.19}$$

respectively.

At this point it is useful to highlight certain properties that all differential equations have in common regardless of their type, order, and whether they are linear or not. First, it is not obvious that just because we can write down a differential equation, a solution exists. The *existence* of a solution to a class of differential equations constitutes an important aspect of the theory of differential equations. Because we are interested in differential equations that arise from applications, their solution should exist. In Section 1.2 we address this question further.

Quite often a differential equation has the solution $y = 0$, a *trivial* solution. For example, if $f(x) = 0$ in Equation 1.1.13, a quick check shows that $y = 0$ is a solution. Trivial solutions are generally of little value.

Another important question is how many solutions does a differential equation have? In physical applications *uniqueness* is not important because, if we are lucky enough to actually find a solution, then its ties to a physical problem usually suggest uniqueness. Nevertheless, the question of uniqueness is of considerable importance in the theory of differential equations. Uniqueness should not be confused with the fact that many solutions to ordinary differential equations contain arbitrary constants, much as indefinite integrals in integral calculus. A solution to a differential equation that has no arbitrary constants is called a *particular solution*.

• Example 1.1.1

Consider the differential equation

$$\frac{dy}{dx} = x + 1, \qquad y(1) = 2. \tag{1.1.20}$$

This condition $y(1) = 2$ is called an *initial condition* and the differential equation plus the initial condition constitute an *initial-value problem*. Straightforward integration yields

$$y(x) = \int (x+1)\,dx + C = \tfrac{1}{2}x^2 + x + C. \tag{1.1.21}$$

Equation 1.1.21 is the *general solution* to the differential equation, Equation 1.1.20, because it is a solution to the differential equation for *every* choice of C. However, if we now satisfy the initial condition $y(1) = 2$, we obtain a *particular solution*. This is done by substituting the corresponding values of x and y into Equation 1.1.21, or

$$2 = \tfrac{1}{2}(1)^2 + 1 + C = \tfrac{3}{2} + C, \quad \text{or} \quad C = \tfrac{1}{2}. \tag{1.1.22}$$

Therefore, the solution to the initial-value problem Equation 1.1.20 is the particular solution

$$y(x) = (x+1)^2/2. \tag{1.1.23}$$

□

Finally, it must be admitted that most differential equations encountered in the "real" world cannot be written down either explicitly or implicitly. For example, the simple differential equation $y' = f(x)$ does not have an analytic solution unless you can integrate $f(x)$. This begs the question of why it is useful to learn analytic techniques for solving differential equations that often fail us. The answer lies in the fact that differential equations that we can solve share many of the same properties and characteristics of differential equations which we can only solve numerically. Therefore, by working with and examining the differential equations that we can solve exactly, we develop our intuition and understanding about those that we can only solve numerically.

Problems

Find the order and state whether the following ordinary differential equations are linear or nonlinear:

1. $y'/y = x^2 + x$
2. $y^2 y' = x + 3$
3. $\sin(y') = 5y$
4. $y''' = y$
5. $y'' = 3x^2$
6. $(y^3)' = 1 - 3y$
7. $y''' = y^3$
8. $y'' - 4y' + 5y = \sin(x)$
9. $y'' + xy = \cos(y'')$
10. $(2x + y)\,dx + (x - 3y)\,dy = 0$
11. $(1 + x^2)y' = (1 + y)^2$
12. $yy'' = x(y^2 + 1)$
13. $y' + y + y^2 = x + e^x$
14. $y''' + \cos(x)y' + y = 0$
15. $x^2 y'' + x^{1/2}(y')^3 + y = e^x$
16. $y''' + xy'' + e^y = x^2$

1.2 SEPARATION OF VARIABLES

The simplest method of solving a first-order ordinary differential equation, if it works, is *separation of variables*. It has the advantage of handling both linear and nonlinear problems, especially *autonomous equations*.[1] From integral calculus, we already met this technique when we solved the first-order differential equation

$$\frac{dy}{dx} = f(x). \tag{1.2.1}$$

[1] An autonomous equation is a differential equation where the independent variable does not explicitly appear in the equation, such as $y' = f(y)$.

By multiplying both sides of Equation 1.2.1 by dx, we obtain

$$dy = f(x)\,dx. \tag{1.2.2}$$

At this point we note that the left side of Equation 1.2.2 contains only y while the right side is purely a function of x. Hence, we can integrate directly and find that

$$y = \int f(x)\,dx + C. \tag{1.2.3}$$

For this technique to work, we must be able to rewrite the differential equation so that all of the y dependence appears on one side of the equation while the x dependence is on the other. Finally we must be able to carry out the integration on both sides of the equation.

One of the interesting aspects of our analysis is the appearance of the arbitrary constant C in Equation 1.2.3. To evaluate this constant we need more information. The most common method is to require that the dependent variable give a particular value for a particular value of x. Because the independent variable x often denotes time, this condition is usually called an *initial condition*, even in cases when the independent variable is not time.

• Example 1.2.1

Let us solve the ordinary differential equation

$$\frac{dy}{dx} = \frac{e^y}{xy}. \tag{1.2.4}$$

Because we can separate variables by rewriting Equation 1.2.4 as

$$ye^{-y}\,dy = \frac{dx}{x}, \tag{1.2.5}$$

its solution is simply

$$-ye^{-y} - e^{-y} = \ln|x| + C \tag{1.2.6}$$

by direct integration. □

• Example 1.2.2

Let us solve

$$\frac{dy}{dx} + y = xe^x y, \tag{1.2.7}$$

subject to the initial condition $y(0) = 1$.

Multiplying Equation 1.2.7 by dx, we find that

$$dy + y\,dx = xe^x y\,dx, \tag{1.2.8}$$

or

$$\frac{dy}{y} = (xe^x - 1)\,dx. \tag{1.2.9}$$

A quick check shows that the left side of Equation 1.2.9 contains only the dependent variable y while the right side depends solely on x and we have separated the variables onto one side or the other. Finally, integrating both sides of this equation, we have

$$\ln(y) = xe^x - e^x - x + C. \tag{1.2.10}$$

Since $y(0) = 1$, $C = 1$ and

$$y(x) = \exp[(x-1)e^x + 1 - x]\,. \tag{1.2.11}$$

In addition to the tried-and-true method of solving ordinary differential equations by hand, scientific computational packages such as MATLAB provide symbolic toolboxes that are designed to do the work for you. In the present case, typing

```
dsolve('Dy+y=x*exp(x)*y','y(0)=1','x')
```

yields

```
ans =
1/exp(-1)*exp(-x+x*exp(x)-exp(x))
```

which is equivalent to Equation 1.2.11.

Our success here should not be overly generalized. Sometimes these toolboxes give the answer in a rather obscure form or they fail completely. For example, in the previous example, MATLAB gives the answer

```
ans =
-lambertw((log(x)+C1)*exp(-1))-1
```

The MATLAB function `lambertw` is Lambert's W function, where `w = lambertw(x)` is the solution to $we^w = x$. Using this definition, we can construct the solution as expressed in Equation 1.2.6. □

• Example 1.2.3

Consider the nonlinear differential equation

$$x^2y' + y^2 = 0. \tag{1.2.12}$$

Separating variables, we find that

$$-\frac{dy}{y^2} = \frac{dx}{x^2}, \quad \text{or} \quad \frac{1}{y} = -\frac{1}{x} + C, \quad \text{or} \quad y = \frac{x}{Cx-1}. \tag{1.2.13}$$

Equation 1.2.13 shows the wide variety of solutions possible for an ordinary differential equation. For example, if we require that $y(0) = 0$, then there are infinitely many different solutions satisfying this initial condition because C can take on any value. On the other hand, if we require that $y(0) = 1$, there is no solution because we cannot choose *any* constant C such that $y(0) = 1$. Finally, if we have the initial condition that $y(1) = 2$, then there is only one possible solution corresponding to $C = \frac{3}{2}$.

Consider now the trial solution $y = 0$. Does it satisfy Equation 1.2.12? Yes, it does. On the other hand, there is no choice of C that yields this solution. The solution $y = 0$ is called a *singular solution* to this equation. *Singular solutions* are solutions to a differential equation that cannot be obtained from a solution with arbitrary constants.

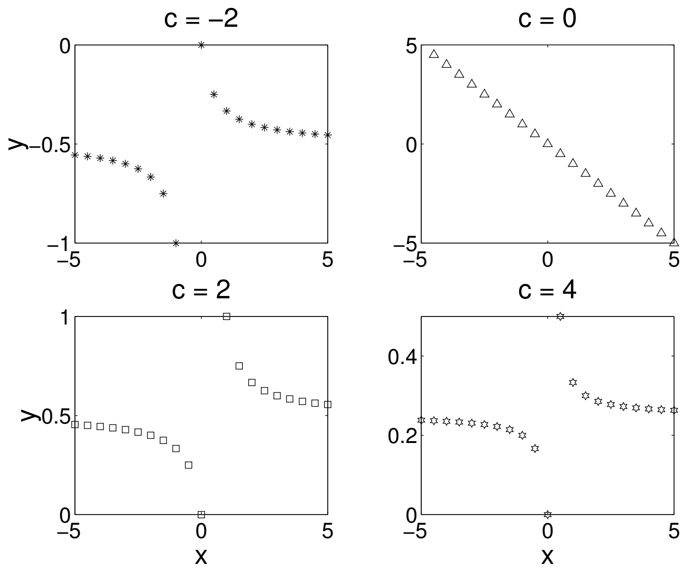


Figure 1.2.1: The solution to Equation 1.2.13 when $C = -2, 0, 2, 4$.

Finally, we illustrate Equation 1.2.13 using MATLAB. This is one of MATLAB's strengths — the ability to convert an abstract equation into a concrete picture. Here the MATLAB script

```
clear
hold on
x = -5:0.5:5;
for c = -2:2:4
  y = x ./ (c*x-1);
  if (c== -2) subplot(2,2,1), plot(x,y,'*')
    axis tight; title('c = -2'); ylabel('y','Fontsize',20); end
  if (c== 0) subplot(2,2,2), plot(x,y,'^')
    axis tight; title('c = 0'); end
  if (c== 2) subplot(2,2,3), plot(x,y,'s')
    axis tight; title('c = 2'); xlabel('x','Fontsize',20);
    ylabel('y','Fontsize',20); end
  if (c== 4) subplot(2,2,4), plot(x,y,'h')
    axis tight; title('c = 4'); xlabel('x','Fontsize',20); end
end
```

yields Figure 1.2.1, which illustrates Equation 1.2.13 when $C = -2, 0, 2$, and 4. □

The previous example showed that first-order ordinary differential equations may have a unique solution, no solution, or many solutions. From a complete study[2] of these equations, we have the following theorem:

---

[2] The proof of the existence and uniqueness of first-order ordinary differential equations is beyond the scope of this book. See Ince, E. L., 1956: *Ordinary Differential Equations*. Dover Publications, Inc., Chapter 3.

Theorem: Existence and Uniqueness

Suppose some real-valued function $f(x,y)$ is continuous on some rectangle in the xy-plane containing the point (a,b) in its interior. Then the initial-value problem

$$\frac{dy}{dx} = f(x,y), \qquad y(a) = b, \tag{1.2.14}$$

has at least one solution on the same open interval I containing the point $x = a$. Furthermore, if the partial derivative $\partial f/\partial y$ is continuous on that rectangle, then the solution is unique on some (perhaps smaller) open interval I_0 containing the point $x = a$. □

• Example 1.2.4

Consider the initial-value problem $y' = 3y^{1/3}/2$ with $y(0) = 1$. Here $f(x,y) = 3y^{1/3}/2$ and $f_y = y^{-2/3}/2$. Because f_y is continuous over a small rectangle containing the point $(0,1)$, there is a unique solution around $x = 0$, namely $y = (x+1)^{3/2}$, which satisfies the differential equation and the initial condition. On the other hand, if the initial condition reads $y(0) = 0$, then f_y is *not* continuous on *any* rectangle containing the point $(0,0)$ and there is no unique solution. For example, two solutions to this initial-value problem, valid on any open interval that includes $x = 0$, are $y_1(x) = x^{3/2}$ and

$$y_2(x) = \begin{cases} (x-1)^{3/2}, & x \ge 1, \\ 0, & x < 1. \end{cases} \tag{1.2.15}$$

□

• Example 1.2.5: Hydrostatic equation

Consider an atmosphere where its density varies only in the vertical direction. The pressure at the surface equals the weight per unit horizontal area of all of the air from sea level to outer space. As you move upward, the amount of air remaining above decreases and so does the pressure. This is why we experience pressure sensations in our ears when ascending or descending in an elevator or airplane. If we rise the small distance dz, there must be a corresponding small decrease in the pressure, dp. This pressure drop must equal the loss of weight in the column per unit area, $-\rho g\, dz$. Therefore, the pressure is governed by the differential equation

$$dp = -\rho g\, dz, \tag{1.2.16}$$

commonly called the *hydrostatic equation*.

To solve Equation 1.2.16, we must express ρ in terms of pressure. For example, in an isothermal atmosphere at constant temperature T_s, the ideal gas law gives $p = \rho R T_s$, where R is the gas constant. Substituting this relationship into our differential equation and separating variables yields

$$\frac{dp}{p} = -\frac{g}{RT_s}\, dz. \tag{1.2.17}$$

Integrating Equation 1.2.17 gives

$$p(z) = p(0) \exp\left(-\frac{gz}{RT_s}\right). \tag{1.2.18}$$

Thus, the pressure (and density) of an isothermal atmosphere decreases exponentially with height. In particular, it decreases by e^{-1} over the distance RT_s/g, the so-called "scale height." □

• Example 1.2.6: Terminal velocity

As an object moves through a fluid, its viscosity resists the motion. Let us find the motion of a mass m as it falls toward the earth under the force of gravity when the drag varies as the square of the velocity.

From Newton's second law, the equation of motion is

$$m\frac{dv}{dt} = mg - C_D v^2, \tag{1.2.19}$$

where v denotes the velocity, g is the gravitational acceleration, and C_D is the drag coefficient. We choose the coordinate system so that a downward velocity is positive.

Equation 1.2.19 can be solved using the technique of separation of variables if we change from time t as the independent variable to the distance traveled x from the point of release. This modification yields the differential equation

$$mv\frac{dv}{dx} = mg - C_D v^2, \tag{1.2.20}$$

since $v = dx/dt$. Separating the variables leads to

$$\frac{v\,dv}{1 - kv^2/g} = g\,dx, \tag{1.2.21}$$

or

$$\ln\left(1 - \frac{kv^2}{g}\right) = -2kx, \tag{1.2.22}$$

where $k = C_D/m$ and $v = 0$ for $x = 0$. Taking the inverse of the natural logarithm, we finally obtain

$$v^2(x) = \frac{g}{k}\left(1 - e^{-2kx}\right). \tag{1.2.23}$$

Thus, as the distance that the object falls increases, so does the velocity, and it eventually approaches a constant value $\sqrt{g/k}$, commonly known as the *terminal velocity*.

Because the drag coefficient C_D varies with the superficial area of the object while the mass depends on the volume, k increases as an object becomes smaller, resulting in a smaller terminal velocity. Consequently, although a human being of normal size will acquire a terminal velocity of approximately 120 mph, a mouse, on the other hand, can fall any distance without injury. □

• Example 1.2.7: Interest rate

Consider a bank account that has been set up to pay out a constant rate of P dollars per year for the purchase of a car. This account has the special feature that it pays an annual interest rate of r on the current balance. We would like to know the balance in the account at any time t.

Although financial transactions occur at regularly spaced intervals, an excellent approximation can be obtained by treating the amount in the account $x(t)$ as a continuous function of time governed by the equation

$$x(t + \Delta t) \approx x(t) + rx(t)\Delta t - P\Delta t, \tag{1.2.24}$$

where we have assumed that both the payment and interest are paid in time increments of Δt. As the time between payments tends to zero, we obtain the first-order ordinary differential equation

$$\frac{dx}{dt} = rx - P. \tag{1.2.25}$$

If we denote the initial deposit into this account by $x(0)$, then at any subsequent time

$$x(t) = x(0)e^{rt} - P\left(e^{rt} - 1\right)/r. \tag{1.2.26}$$

Although we could compute $x(t)$ as a function of P, r, and $x(0)$, there are only three separate cases that merit our close attention. If $P/r > x(0)$, then the account will eventually equal zero at $rt = \ln\{P/\,[P - rx(0)]\}$. On the other hand, if $P/r < x(0)$, the amount of money in the account will grow without bound. Finally, the case $x(0) = P/r$ is the equilibrium case where the amount of money paid out balances the growth of money due to interest so that the account always has the balance of P/r. □

• Example 1.2.8: Steady-state flow of heat

When the inner and outer walls of a body, for example the inner and outer walls of a house, are maintained at *different constant* temperatures, heat will flow from the warmer wall to the colder one. When each surface parallel to a wall has attained a constant temperature, the flow of heat has reached a steady state. In a steady-state flow of heat, each surface parallel to a wall, because its temperature is now constant, is referred to as an isothermal surface. Isothermal surfaces at different distances from an interior wall will have different temperatures. In many cases the temperature of an isothermal surface is only a function of its distance x from the interior wall, and the rate of flow of heat Q in a unit time across such a surface is proportional both to the area A of the surface and to dT/dx, where T is the temperature of the isothermal surface. Hence,

$$Q = -\kappa A\frac{dT}{dx}, \tag{1.2.27}$$

where κ is called the thermal conductivity of the material between the walls.

In place of a flat wall, let us consider a hollow cylinder whose inner and outer surfaces are located at $r = r_1$ and $r = r_2$, respectively. At steady state, Equation 1.2.27 becomes

$$Q_r = -\kappa A\frac{dT}{dr} = -\kappa(2\pi rL)\frac{dT}{dr}, \tag{1.2.28}$$

assuming no heat generation within the cylindrical wall.

We can find the temperature distribution inside the cylinder by solving Equation 1.2.28 along with the appropriate conditions on $T(r)$ at $r = r_1$ and $r = r_2$ (the boundary conditions). To illustrate the wide choice of possible boundary conditions, let us require that inner surface is maintained at the temperature T_1. We assume that along the outer surface heat is lost by convection to the environment, which has the temperature T_∞. This heat loss is usually modeled by the equation

$$\kappa \left.\frac{dT}{dr}\right|_{r=r_2} = -h(T - T_\infty), \tag{1.2.29}$$

where $h > 0$ is the convective heat transfer coefficient. Upon integrating Equation 1.2.28,

$$T(r) = -\frac{Q_r}{2\pi\kappa L}\ln(r) + C, \tag{1.2.30}$$

where Q_r is also an unknown. Substituting Equation 1.2.30 into the boundary conditions, we obtain

$$T(r) = T_1 + \frac{Q_r}{2\pi\kappa L}\ln(r_1/r), \tag{1.2.31}$$

with

$$Q_r = \frac{2\pi\kappa L(T_1 - T_\infty)}{\kappa/r_2 + h\ln(r_2/r_1)}. \tag{1.2.32}$$

As r_2 increases, the first term in the denominator of Equation 1.2.32 decreases while the second term increases. Therefore, Q_r has its largest magnitude when the denominator is smallest, assuming a fixed numerator. This occurs at the critical radius $r_{cr} = \kappa/h$, where

$$Q_r^{max} = \frac{2\pi\kappa L(T_1 - T_\infty)}{1 + \ln(r_{cr}/r_1)}. \tag{1.2.33}$$

□

• **Example 1.2.9: Population dynamics**

Consider a population $P(t)$ that can change only by a birth or death but not by immigration or emigration. If $B(t)$ and $D(t)$ denote the number of births or deaths, respectively, as a function of time t, the *birth rate* and *death rate* (in births or deaths per unit time) is

$$b(t) = \lim_{\Delta t\to 0}\frac{B(t+\Delta t) - B(t)}{P(t)\Delta t} = \frac{1}{P}\frac{dB}{dt}, \tag{1.2.34}$$

and

$$d(t) = \lim_{\Delta t\to 0}\frac{D(t+\Delta t) - D(t)}{P(t)\Delta t} = \frac{1}{P}\frac{dD}{dt}. \tag{1.2.35}$$

Now,

$$P'(t) = \lim_{\Delta t\to 0}\frac{P(t+\Delta t) - P(t)}{\Delta t} \tag{1.2.36}$$

$$= \lim_{\Delta t\to 0}\frac{[B(t+\Delta t) - B(t)] - [D(t+\Delta t) - D(t)]}{\Delta t} \tag{1.2.37}$$

$$= B'(t) - D'(t). \tag{1.2.38}$$

Therefore,

$$P'(t) = [b(t) - d(t)]P(t). \tag{1.2.39}$$

When the birth and death rates are constants, namely $\bar{b}$ and $\bar{d}$, respectively, the population evolves according to

$$P(t) = P(0) \exp\left[\left(\bar{b} - \bar{d}\right) t\right]. \tag{1.2.40}$$

□

• Example 1.2.10: Logistic equation

The study of population dynamics yields an important class of first-order, nonlinear, ordinary differential equations: the logistic equation. This equation arose in Pierre François Verhulst's (1804–1849) study of animal populations.[3] If $x(t)$ denotes the number of species in the population and k is the (constant) environment capacity (the number of species that can simultaneously live in the geographical region), then the logistic or Verhulst's equation is

$$x' = ax(k - x)/k, \tag{1.2.41}$$

where a is the population growth rate for a small number of species.

To solve Equation 1.2.41, we rewrite it as

$$\frac{dx}{(1 - x/k)x} = \frac{dx}{x} + \frac{x/k}{1 - x/k}\,dx = r\,dt. \tag{1.2.42}$$

Integration yields

$$\ln|x| - \ln|1 - x/k| = rt + \ln(C), \tag{1.2.43}$$

or

$$\frac{x}{1 - x/k} = Ce^{rt}. \tag{1.2.44}$$

If $x(0) = x_0$,

$$x(t) = \frac{kx_0}{x_0 + (k - x_0)e^{-rt}}. \tag{1.2.45}$$

As $t \to \infty$, $x(t) \to k$, the asymptotically stable solution. □

• Example 1.2.11: Chemical reactions

Chemical reactions are often governed by first-order ordinary differential equations. For example, first-order reactions, which describe reactions of the form $\mathrm{A} \xrightarrow{k} \mathrm{B}$, yield the differential equation

$$-\frac{1}{a}\frac{d[\mathrm{A}]}{dt} = k[\mathrm{A}], \tag{1.2.46}$$

where k is the rate at which the reaction is taking place. Because for every molecule of A that disappears one molecule of B is produced, $a = 1$ and Equation 1.2.46 becomes

$$-\frac{d[\mathrm{A}]}{dt} = k[\mathrm{A}]. \tag{1.2.47}$$

---

[3] Verhulst, P. F., 1838: Notice sur la loi que la population suit dans son accroissement. *Correspond. Math. Phys.*, **10**, 113–121.

Integration of Equation 1.2.47 leads to

$$-\int \frac{d[\mathrm{A}]}{[\mathrm{A}]} = k \int dt. \qquad (\mathbf{1.2.48})$$

If we denote the initial value of [A] by $[\mathrm{A}]_0$, then integration yields

$$-\ln{[\mathrm{A}]} = kt - \ln{[\mathrm{A}]_0}, \qquad (\mathbf{1.2.49})$$

or

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-kt}. \qquad (\mathbf{1.2.50})$$

The exponential form of the solution suggests that there is a *time constant* τ, which is called the *decay time* of the reaction. This quantity gives the time required for the concentration of decrease by $1/e$ of its initial value $[\mathrm{A}]_0$. It is given by $\tau = 1/k$.

Turning to second-order reactions, there are two cases. The first is a reaction between two identical species: $\mathrm{A} + \mathrm{A} \stackrel{k}{\rightarrow}$ products. The rate expression here is

$$-\frac{1}{2}\frac{d[\mathrm{A}]}{dt} = k[\mathrm{A}]^2. \qquad (\mathbf{1.2.51})$$

The second case is an overall second-order reaction between two unlike species, given by $\mathrm{A} + \mathrm{B} \stackrel{k}{\rightarrow} \mathrm{X}$. In this case, the reaction is first order in each of the reactants A and B and the rate expression is

$$-\frac{d[\mathrm{A}]}{dt} = k[\mathrm{A}][\mathrm{B}]. \qquad (\mathbf{1.2.52})$$

Turning to Equation 1.2.51 first, we have by separation of variables

$$-\int_{[\mathrm{A}]_0}^{[\mathrm{A}]} \frac{d[\mathrm{A}]}{[\mathrm{A}]^2} = 2k \int_0^t d\tau, \qquad (\mathbf{1.2.53})$$

or

$$\frac{1}{[\mathrm{A}]} = \frac{1}{[\mathrm{A}]_0} + 2kt. \qquad (\mathbf{1.2.54})$$

Therefore, a plot of the inverse of A versus time will yield a straight line with slope equal to $2k$ and intercept $1/[\mathrm{A}]_0$.

With regard to Equation 1.2.52, because an increase in X must be at the expense of A and B, it is useful to express the rate equation in terms of the concentration of X, $[\mathrm{X}] = [\mathrm{A}]_0 - [\mathrm{A}] = [\mathrm{B}]_0 - [\mathrm{B}]$, where $[\mathrm{A}]_0$ and $[\mathrm{B}]_0$ are the initial concentrations. Then, this equation becomes

$$\frac{d[\mathrm{X}]}{dt} = k\left([\mathrm{A}]_0 - [\mathrm{X}]\right)\left([\mathrm{B}]_0 - [\mathrm{X}]\right). \qquad (\mathbf{1.2.55})$$

Separation of variables leads to

$$\int_{[\mathrm{X}]_0}^{[\mathrm{X}]} \frac{d\xi}{\left([\mathrm{A}]_0 - \xi\right)\left([\mathrm{B}]_0 - \xi\right)} = k \int_0^t d\tau. \qquad (\mathbf{1.2.56})$$

To integrate the left side, we rewrite the integral

$$\int \frac{d\xi}{\left([\mathrm{A}]_0 - \xi\right)\left([\mathrm{B}]_0 - \xi\right)} = \int \frac{d\xi}{\left([\mathrm{A}]_0 - [\mathrm{B}]_0\right)\left([\mathrm{B}]_0 - \xi\right)} - \int \frac{d\xi}{\left([\mathrm{A}]_0 - [\mathrm{B}]_0\right)\left([\mathrm{A}]_0 - \xi\right)}. \qquad (\mathbf{1.2.57})$$

Carrying out the integration,

$$\frac{1}{[\mathrm{A}]_0 - [\mathrm{B}]_0} \ln\left(\frac{[\mathrm{B}]_0[\mathrm{A}]}{[\mathrm{A}]_0[\mathrm{B}]}\right) = kt. \tag{1.2.58}$$

Again the reaction rate constant k can be found by plotting the data in the form of the left side of Equation 1.2.58 against t.

Problems

For Problems 1–10, solve the following ordinary differential equations by separation of variables. Then use MATLAB to plot your solution. Try and find the symbolic solution using MATLAB's `dsolve`.

1. $\dfrac{dy}{dx} = xe^y$

2. $(1+y^2)\,dx - (1+x^2)\,dy = 0$

3. $\ln(x)\dfrac{dx}{dy} = xy$

4. $\dfrac{y^2}{x}\dfrac{dy}{dx} = 1 + x^2$

5. $\dfrac{dy}{dx} = \dfrac{2x + xy^2}{y + x^2 y}$

6. $\dfrac{dy}{dx} = (xy)^{1/3}$

7. $\dfrac{dy}{dx} = e^{x+y}$

8. $\dfrac{dy}{dx} = (x^3+5)(y^2+1)$

9. Solve the initial-value problem

$$\frac{dy}{dt} = -ay + \frac{b}{y^2}, \qquad y(0) = y_0,$$

where a and b are constants.

10. Setting $u = y - x$, solve the first-order ordinary differential equation

$$\frac{dy}{dx} = \frac{y-x}{x^2} + 1.$$

11. Using the hydrostatic equation, show that the pressure within an atmosphere where the temperature decreases uniformly with height, $T(z) = T_0 - \Gamma z$, varies as

$$p(z) = p_0 \left(\frac{T_0 - \Gamma z}{T_0}\right)^{g/(R\Gamma)},$$

where p_0 is the pressure at $z = 0$.

12. Using the hydrostatic equation, show that the pressure within an atmosphere with the temperature distribution

$$T(z) = \begin{cases} T_0 - \Gamma z, & 0 \le z \le H, \\ T_0 - \Gamma H, & H \le z, \end{cases}$$

is

$$p(z) = p_0 \begin{cases} \left(\dfrac{T_0 - \Gamma z}{T_0}\right)^{g/(R\Gamma)}, & 0 \leq z \leq H, \\ \left(\dfrac{T_0 - \Gamma H}{T_0}\right)^{g/(R\Gamma)} \exp\left[-\dfrac{g(z-H)}{R(T_0 - \Gamma H)}\right], & H \leq z, \end{cases}$$

where p_0 is the pressure at $z = 0$.

13. The voltage V as a function of time t within an electrical circuit[4] consisting of a capacitor with capacitance C and a diode in series is governed by the first-order ordinary differential equation

$$C\frac{dV}{dt} + \frac{V}{R} + \frac{V^2}{S} = 0,$$

where R and S are positive constants. If the circuit initially has a voltage V_0 at $t = 0$, find the voltage at subsequent times.

14. A glow plug is an electrical element inside a reaction chamber, which either ignites the nearby fuel or warms the air in the chamber so that the ignition will occur more quickly. An accurate prediction of the wire's temperature is important in the design of the chamber.

Assuming that heat convection and conduction are not important,[5] the temperature T of the wire is governed by

$$A\frac{dT}{dt} + B(T^4 - T_a^4) = P,$$

where A equals the specific heat of the wire times its mass, B equals the product of the emissivity of the surrounding fluid times the wire's surface area times the Stefan-Boltzmann constant, T_a is the temperature of the surrounding fluid, and P is the power input. The temperature increases due to electrical resistance and is reduced by radiation to the surrounding fluid.

Show that the temperature is given by

$$\frac{4B\gamma^3 t}{A} = 2\left[\tan^{-1}\left(\frac{T}{\gamma}\right) - \tan^{-1}\left(\frac{T_0}{\gamma}\right)\right] - \ln\left[\frac{(T-\gamma)(T_0+\gamma)}{(T+\gamma)(T_0-\gamma)}\right],$$

where $\gamma^4 = P/B + T_a^4$ and T_0 is the initial temperature of the wire.

15. Let us denote the number of tumor cells by $N(t)$. Then a widely used deterministic tumor growth law[6] is

$$\frac{dN}{dt} = bN\ln(K/N),$$

where K is the largest tumor size and $1/b$ is the length of time required for the specific growth to decrease by $1/e$. If the initial value of $N(t)$ is $N(0)$, find $N(t)$ at any subsequent time t.

---

[4] See Aiken, C. B., 1938: Theory of the diode voltmeter. *Proc. IRE*, **26**, 859–876.

[5] See Clark, S. K., 1956: Heat-up time of wire glow plugs. *Jet Propulsion*, **26**, 278–279.

[6] See Hanson, F. B., and C. Tier, 1982: A stochastic model of tumor growth. *Math. Biosci.*, **61**, 73–100.

16. The drop in laser intensity in the direction of propagation x due to one- and two-photon absorption in photosensitive glass is governed[7] by

$$\frac{dI}{dx} = -\alpha I - \beta I^2,$$

where I is the laser intensity, α and β are the single-photon and two-photon coefficients, respectively. Show that the laser intensity distribution is

$$I(x) = \frac{\alpha I(0) e^{-\alpha x}}{\alpha + \beta I(0)\left(1 - e^{-\alpha x}\right)},$$

where $I(0)$ is the laser intensity at the entry point of the media, $x = 0$.

17. The third-order reaction $\mathrm{A} + \mathrm{B} + \mathrm{C} \stackrel{k}{\rightarrow} \mathrm{X}$ is governed by the kinetics equation

$$\frac{d[\mathrm{X}]}{dt} = k\left([\mathrm{A}]_0 - [\mathrm{X}]\right)\left([\mathrm{B}]_0 - [\mathrm{X}]\right)\left([\mathrm{C}]_0 - [\mathrm{X}]\right),$$

where $[\mathrm{A}]_0$, $[\mathrm{B}]_0$, and $[\mathrm{C}]_0$ denote the initial concentration of A, B, and C, respectively. Find how $[\mathrm{X}]$ varies with time t.

18. The reversible reaction $\mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{B}$ is described by the kinetics equation[8]

$$\frac{d[\mathrm{X}]}{dt} = k_1\left([\mathrm{A}]_0 - [\mathrm{X}]\right) - k_2\left([\mathrm{B}]_0 + [\mathrm{X}]\right),$$

where $[\mathrm{X}]$ denotes the increase in the concentration of B while $[\mathrm{A}]_0$ and $[\mathrm{B}]_0$ are the initial concentrations of A and B, respectively. Find $[\mathrm{X}]$ as a function of time t. Hint: Show that this differential equation can be written

$$\frac{d[\mathrm{X}]}{dt} = (k_1 - k_2)\left(\alpha + [\mathrm{X}]\right), \qquad \alpha = \frac{k_1[\mathrm{A}]_0 - k_2[\mathrm{B}]_0}{k_1 + k_2}.$$

1.3 HOMOGENEOUS EQUATIONS

A *homogeneous ordinary differential equation* is a differential equation of the form

$$M(x,y)\,dx + N(x,y)\,dy = 0, \qquad \textbf{(1.3.1)}$$

where both $M(x,y)$ and $N(x,y)$ are homogeneous functions of the same degree n. That means: $M(tx,ty) = t^n M(x,y)$ and $N(tx,ty) = t^n N(x,y)$. For example, the ordinary differential equation

$$(x^2 + y^2)\,dx + (x^2 - xy)\,dy = 0 \qquad \textbf{(1.3.2)}$$

---

[7] See Weitzman, P. S., and U. Österberg, 1996: Two-photon absorption and photoconductivity in photosensitive glasses. *J. Appl. Phys.*, **79**, 8648–8655.

[8] See Küster, F. W., 1895: Ueber den Verlauf einer umkehrbaren Reaktion erster Ordnung in homogenem System. *Z. Physik. Chem.*, **18**, 171–179.

is a homogeneous equation because both coefficients are homogeneous functions of degree 2:

$$M(tx,ty) = t^2x^2 + t^2y^2 = t^2(x^2+y^2) = t^2M(x,y), \tag{1.3.3}$$

and

$$N(tx,ty) = t^2x^2 - t^2xy = t^2(x^2-xy) = t^2N(x,y). \tag{1.3.4}$$

Why is it useful to recognize homogeneous ordinary differential equations? Let us set $y = ux$ so that Equation 1.3.2 becomes

$$(x^2 + u^2x^2)\,dx + (x^2 - ux^2)(u\,dx + x\,du) = 0. \tag{1.3.5}$$

Then,

$$x^2(1+u)\,dx + x^3(1-u)\,du = 0, \tag{1.3.6}$$

$$\frac{1-u}{1+u}\,du + \frac{dx}{x} = 0, \tag{1.3.7}$$

or

$$\left(-1 + \frac{2}{1+u}\right)du + \frac{dx}{x} = 0. \tag{1.3.8}$$

Integrating Equation 1.3.8,

$$-u + 2\ln|1+u| + \ln|x| = \ln|c|, \tag{1.3.9}$$

$$-\frac{y}{x} + 2\ln\left|1 + \frac{y}{x}\right| + \ln|x| = \ln|c|, \tag{1.3.10}$$

$$\ln\left[\frac{(x+y)^2}{cx}\right] = \frac{y}{x}, \tag{1.3.11}$$

or

$$(x+y)^2 = cxe^{y/x}. \tag{1.3.12}$$

Problems

First show that the following differential equations are homogeneous and then find their solution. Then use MATLAB to plot your solution. Try and find the symbolic solution using MATLAB's `dsolve`.

1. $(x+y)\dfrac{dy}{dx} = y$
2. $(x+y)\dfrac{dy}{dx} = x - y$
3. $2xy\dfrac{dy}{dx} = -(x^2+y^2)$
4. $x(x+y)\dfrac{dy}{dx} = y(x-y)$
5. $xy' = y + 2\sqrt{xy}$
6. $xy' = y - \sqrt{x^2+y^2}$
7. $y' = \sec(y/x) + y/x$
8. $y' = e^{y/x} + y/x.$

1.4 EXACT EQUATIONS

Consider the multivariable function $z = f(x,y)$. Then the total derivative is

$$dz = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy = M(x,y)\,dx + N(x,y)\,dy. \tag{1.4.1}$$

If the solution to a first-order ordinary differential equation can be written as $f(x, y) = c$, then the corresponding differential equation is

$$M(x, y)\, dx + N(x, y)\, dy = 0. \tag{1.4.2}$$

How do we know if we have an *exact equation*, Equation 1.4.2? From the definition of $M(x, y)$ and $N(x, y)$,

$$\frac{\partial M}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial N}{\partial x}, \tag{1.4.3}$$

if $M(x, y)$ and $N(x, y)$ and their first-order partial derivatives are continuous. Consequently, if we can show that our ordinary differential equation is exact, we can integrate

$$\frac{\partial f}{\partial x} = M(x, y) \qquad \text{and} \qquad \frac{\partial f}{\partial y} = N(x, y) \tag{1.4.4}$$

to find the solution $f(x, y) = c$.

• Example 1.4.1

Let us check and see if

$$[y^2 \cos(x) - 3x^2 y - 2x]\, dx + [2y \sin(x) - x^3 + \ln(y)]\, dy = 0 \tag{1.4.5}$$

is exact.

Since $M(x, y) = y^2 \cos(x) - 3x^2 y - 2x$, and $N(x, y) = 2y \sin(x) - x^3 + \ln(y)$, we find that

$$\frac{\partial M}{\partial y} = 2y \cos(x) - 3x^2, \tag{1.4.6}$$

and

$$\frac{\partial N}{\partial x} = 2y \cos(x) - 3x^2. \tag{1.4.7}$$

Because $N_x = M_y$, Equation 1.4.5 is an exact equation. □

• Example 1.4.2

Because Equation 1.4.5 is an exact equation, let us find its solution. Starting with

$$\frac{\partial f}{\partial x} = M(x, y) = y^2 \cos(x) - 3x^2 y - 2x, \tag{1.4.8}$$

direct integration gives

$$f(x, y) = y^2 \sin(x) - x^3 y - x^2 + g(y). \tag{1.4.9}$$

Substituting Equation 1.4.9 into the equation $f_y = N$, we obtain

$$\frac{\partial f}{\partial y} = 2y \sin(x) - x^3 + g'(y) = 2y \sin(x) - x^3 + \ln(y). \tag{1.4.10}$$

Thus, $g'(y) = \ln(y)$, or $g(y) = y \ln(y) - y + C$. Therefore, the solution to the ordinary differential equation, Equation 1.4.5, is

$$y^2 \sin(x) - x^3 y - x^2 + y \ln(y) - y = c. \tag{1.4.11}$$

□

• **Example 1.4.3**

Consider the differential equation

$$(x + y)\,dx + x\ln(x)\,dy = 0 \tag{1.4.12}$$

on the interval $(0, \infty)$. A quick check shows that Equation 1.4.12 is not exact since

$$\frac{\partial M}{\partial y} = 1, \qquad \text{and} \qquad \frac{\partial N}{\partial x} = 1 + \ln(x). \tag{1.4.13}$$

However, if we multiply Equation 1.4.12 by $1/x$ so that it becomes

$$\left(1 + \frac{y}{x}\right)\,dx + \ln(x)\,dy = 0, \tag{1.4.14}$$

then this modified differential equation is exact because

$$\frac{\partial M}{\partial y} = \frac{1}{x}, \qquad \text{and} \qquad \frac{\partial N}{\partial x} = \frac{1}{x}. \tag{1.4.15}$$

Therefore, the solution to Equation 1.4.12 is

$$x + y\ln(x) = C. \tag{1.4.16}$$

This mysterious function that converts an inexact differential equation into an exact one is called an *integrating factor*. Unfortunately there is no general rule for finding one unless the equation is linear.

Problems

Show that the following equations are exact. Then solve them, using MATLAB to plot them. Finally, try and find the symbolic solution using MATLAB's `dsolve`.

1. $2xyy' = x^2 - y^2$
2. $(x + y)y' + y = x$
3. $(y^2 - 1)\,dx + [2xy - \sin(y)]\,dy = 0$
4. $[\sin(y) - 2xy + x^2]\,dx$ $+ [x\cos(y) - x^2]\,dy = 0$
5. $-y\,dx/x^2 + (1/x + 1/y)\,dy = 0$
6. $(3x^2 - 6xy)\,dx - (3x^2 + 2y)\,dy = 0$
7. $y\sin(xy)\,dx + x\sin(xy)\,dy = 0$
8. $(2xy^2 + 3x^2)\,dx + 2x^2y\,dy = 0$
9. $(2xy^3 + 5x^4y)\,dx$ $+ (3x^2y^2 + x^5 + 1)\,dy = 0$
10. $(x^3 + y/x)\,dx + [y^2 + \ln(x)]\,dy = 0$
11. $[x + e^{-y} + x\ln(y)]\,dy$ $+ [y\ln(y) + e^x]\,dx = 0$
12. $\cos(4y^2)\,dx - 8xy\sin(4y^2)\,dy = 0$
13. $\sin^2(x + y)\,dx - \cos^2(x + y)\,dy = 0$

14. Show that the integrating factor for $(x-y)y' + \alpha y(1-y) = 0$ is $\mu(y) = y^a/(1-y)^{a+2}$, $a+1 = 1/\alpha$. Then show that the solution is

$$\alpha x \frac{y^{a+1}}{(1-y)^{a+1}} - \int_0^y \frac{\xi^{a+1}}{(1-\xi)^{a+2}}\, d\xi = C.$$

1.5 LINEAR EQUATIONS

In the case of first-order ordinary differential equations, any differential equation of the form

$$a_1(x)\frac{dy}{dx} + a_0(x)y = f(x) \tag{1.5.1}$$

is said to be linear.

Consider now the linear ordinary differential equation

$$x\frac{dy}{dx} - 4y = x^6 e^x \tag{1.5.2}$$

or

$$\frac{dy}{dx} - \frac{4}{x}y = x^5 e^x. \tag{1.5.3}$$

Let us now multiply Equation 1.5.3 by x^{-4}. (How we knew that it should be x^{-4} and not something else will be addressed shortly.) This magical factor is called an *integrating factor* because Equation 1.5.3 can be rewritten

$$\frac{1}{x^4}\frac{dy}{dx} - \frac{4}{x^5}y = xe^x, \tag{1.5.4}$$

or

$$\frac{d}{dx}\left(\frac{y}{x^4}\right) = xe^x. \tag{1.5.5}$$

Thus, our introduction of the integrating factor x^{-4} allows us to use the differentiation product rule in reverse and collapse the right side of Equation 1.5.4 into a single x derivative of a function of x times y. If we had selected the incorrect integrating factor, the right side would not have collapsed into this useful form.

With Equation 1.5.5, we may integrate both sides and find that

$$\frac{y}{x^4} = \int xe^x\, dx + C, \tag{1.5.6}$$

or

$$\frac{y}{x^4} = (x-1)e^x + C, \tag{1.5.7}$$

or

$$y = x^4(x-1)e^x + Cx^4. \tag{1.5.8}$$

From this example, it is clear that finding the integrating factor is crucial to solving first-order, linear, ordinary differential equations. To do this, let us first rewrite Equation 1.5.1 by dividing through by $a_1(x)$ so that it becomes

$$\frac{dy}{dx} + P(x)y = Q(x), \tag{1.5.9}$$

or

$$dy + [P(x)y - Q(x)]\ dx = 0. \tag{1.5.10}$$

If we denote the integrating factor by $\mu(x)$, then

$$\mu(x)dy + \mu(x)[P(x)y - Q(x)]\, dx = 0. \tag{1.5.11}$$

Clearly, we can solve Equation 1.5.11 by direct integration if it is an exact equation. If this is true, then

$$\frac{\partial \mu}{\partial x} = \frac{\partial}{\partial y}\left\{\mu(x)[P(x)y - Q(x)]\right\}, \tag{1.5.12}$$

or

$$\frac{d\mu}{dx} = \mu(x)P(x), \qquad \text{and} \qquad \frac{d\mu}{\mu} = P(x)\, dx. \tag{1.5.13}$$

Integrating Equation 1.5.13,

$$\mu(x) = \exp\left[\int^x P(\xi)\, d\xi\right]. \tag{1.5.14}$$

Note that we do not need a constant of integration in Equation 1.5.14 because Equation 1.5.11 is unaffected by a constant multiple. It is also interesting that the integrating factor only depends on $P(x)$ and not $Q(x)$.

We can summarize our findings in the following theorem.

Theorem: Linear First-Order Equation

If the functions $P(x)$ and $Q(x)$ are continuous on the open interval I containing the point x_0, then the initial-value problem

$$\frac{dy}{dx} + P(x)y = Q(x), \qquad y(x_0) = y_0,$$

has a unique solution $y(x)$ on I, given by

$$y(x) = \frac{C}{\mu(x)} + \frac{1}{\mu(x)}\int^x Q(\xi)\mu(\xi)\, d\xi$$

with an appropriate value of C, and $\mu(x)$ is defined by Equation 1.5.14. □

The procedure for implementing this theorem is as follows:

- **Step 1:** If necessary, divide the differential equation by the coefficient of dy/dx. This gives an equation of the form Equation 1.5.9 and we can find $P(x)$ by inspection.

- **Step 2:** Find the integrating factor by Equation 1.5.14.

- **Step 3:** Multiply the equation created in Step 1 by the integrating factor.

- **Step 4:** Run the derivative product rule in reverse, collapsing the left side of the differential equation into the form $d[\mu(x)y]/dx$. If you are unable to do this, you have made a mistake.

- **Step 5:** Integrate both sides of the differential equation to find the solution.

The following examples illustrate the technique.

• **Example 1.5.1**

Let us solve the linear, first-order ordinary differential equation

$$xy' - y = 4x\ln(x). \tag{1.5.15}$$

We begin by dividing through by x to convert Equation 1.5.15 into its canonical form. This yields

$$y' - \frac{1}{x}y = 4\ln(x). \tag{1.5.16}$$

From Equation 1.5.16, we see that $P(x) = 1/x$. Consequently, from Equation 1.5.14, we have that

$$\mu(x) = \exp\left[\int^x P(\xi)\,d\xi\right] = \exp\left(-\int^x \frac{d\xi}{\xi}\right) = \frac{1}{x}. \tag{1.5.17}$$

Multiplying Equation 1.5.16 by the integrating factor, we find that

$$\frac{y'}{x} - \frac{y}{x^2} = \frac{4\ln(x)}{x}, \tag{1.5.18}$$

or

$$\frac{d}{dx}\left(\frac{y}{x}\right) = \frac{4\ln(x)}{x}. \tag{1.5.19}$$

Integrating both sides of Equation 1.5.19,

$$\frac{y}{x} = 4\int \frac{\ln(x)}{x}dx = 2\ln^2(x) + C. \tag{1.5.20}$$

Multiplying Equation 1.5.20 through by x yields the general solution

$$y = 2x\ln^2(x) + Cx. \tag{1.5.21}$$

Although it is nice to have a closed-form solution, considerable insight can be gained by graphing the solution for a wide variety of initial conditions. To illustrate this, consider the MATLAB script

```
clear
% use symbolic toolbox to solve Equation 1.5.15
y = dsolve('x*Dy-y=4*x*log(x)','y(1) = c','x');
% take the symbolic version of the solution
%     and convert it into executable code
solution = inline(vectorize(y),'x','c');
close all; axes; hold on
% now plot the solution for a wide variety of initial conditions
x = 0.1:0.1:2;
for c = -2:4
  if (c==-2) plot(x,solution(x,c),'.'); end
  if (c==-1) plot(x,solution(x,c),'o'); end
  if (c== 0) plot(x,solution(x,c),'x'); end
```

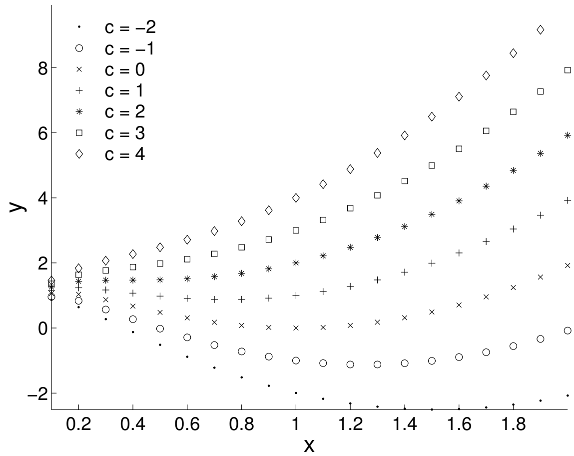


Figure 1.5.1: The solution to Equation 1.5.15 when the initial condition is $y(1) = c$.

```
  if (c== 1) plot(x,solution(x,c),'+'); end
  if (c== 2) plot(x,solution(x,c),'*'); end
  if (c== 3) plot(x,solution(x,c),'s'); end
  if (c== 4) plot(x,solution(x,c),'d'); end
end
axis tight
xlabel('x','Fontsize',20); ylabel('y','Fontsize',20)
legend('c = -2','c = -1','c = 0','c = 1',...
       'c = 2','c = 3','c = 4'); legend boxoff
```

This script does two things. First, it uses MATLAB's symbolic toolbox to solve Equation 1.5.15. Alternatively, we could have used Equation 1.5.21 and introduced it as a function. The second portion of this script plots this solution for $y(1) = C$ where $C = -2, -1, 0, 1, 2, 3, 4$. Figure 1.5.1 shows the results. As $x \to 0$, we note how all of the solutions behave like $2x\ln^2(x)$. □

• Example 1.5.2

Let us solve the first-order ordinary differential equation

$$\frac{dy}{dx} = \frac{y}{y-x} \tag{1.5.22}$$

subject to the initial condition $y(2) = 6$.

Beginning as before, we rewrite Equation 1.5.22 in the canonical form

$$(y-x)y' - y = 0. \tag{1.5.23}$$

Examining Equation 1.5.23 more closely, we see that it is a nonlinear equation in y. On the other hand, if we treat x as the *dependent* variable and y as the *independent variable*, we can write Equation 1.5.23 as the *linear* equation

$$\frac{dx}{dy} + \frac{x}{y} = 1. \tag{1.5.24}$$

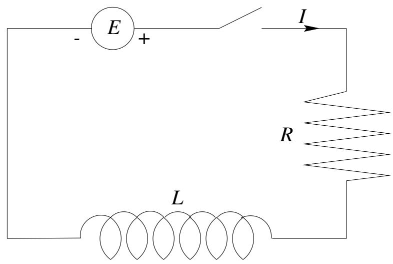


Figure 1.5.2: Schematic diagram for an electric circuit that contains a resistor of resistance R and an inductor of inductance L.

Proceeding as before, we have that $P(y) = 1/y$ and $\mu(y) = y$ so that Equation 1.5.24 can be rewritten

$$\frac{d}{dy}(yx) = y \tag{1.5.25}$$

or

$$yx = \tfrac{1}{2}y^2 + C. \tag{1.5.26}$$

Introducing the initial condition, we find that $C = -6$. Solving for y, we obtain

$$y = x \pm \sqrt{x^2 + 12}. \tag{1.5.27}$$

We must take the positive sign in order that $y(2) = 6$ and

$$y = x + \sqrt{x^2 + 12}. \tag{1.5.28}$$

□

• Example 1.5.3: Electric circuits

A rich source of first-order differential equations is the analysis of simple electrical circuits. These electrical circuits are constructed from three fundamental components: the resistor, the inductor, and the capacitor. Each of these devices gives the following voltage drop: In the case of a resistor, the voltage drop equals the product of the resistance R times the current I. For the inductor, the voltage drop is $L\, dI/dt$, where L is called the inductance, while the voltage drop for a capacitor equals Q/C, where Q is the instantaneous charge and C is called the capacitance.

How are these voltage drops applied to mathematically describe an electrical circuit? This question leads to one of the fundamental laws in physics, **Kirchhoff's law**:*The algebraic sum of all the voltage drops around an electric loop or circuit is zero.*

To illustrate Kirchhoff's law, consider the electrical circuit shown in Figure 1.5.2. By Kirchhoff's law, the electromotive force E, provided by a battery, for example, equals the sum of the voltage drops across the resistor RI and $L\, dI/dt$. Thus the (differential) equation that governs this circuit is

$$L\frac{dI}{dt} + RI = E. \tag{1.5.29}$$

Assuming that E, I, and R are constant, we can rewrite Equation 1.5.29 as

$$\frac{d}{dt}\left[e^{Rt/L}I(t)\right] = \frac{E}{L}e^{Rt/L}. \tag{1.5.30}$$

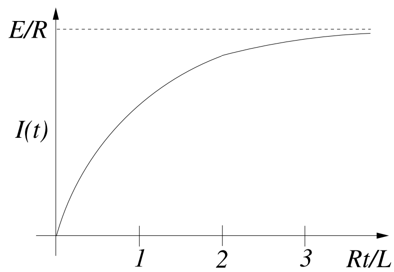


Figure 1.5.3: The temporal evolution of current $I(t)$ inside an electrical circuit shown in Figure 1.5.2 with a constant electromotive force E.

Integrating both sides of Equation 1.5.30,

$$e^{Rt/L} I(t) = \frac{E}{R} e^{Rt/L} + C_1, \tag{1.5.31}$$

or

$$I(t) = \frac{E}{R} + C_1 e^{-Rt/L}. \tag{1.5.32}$$

To determine C_1, we apply the initial condition. Because the circuit is initially dead, $I(0) = 0$, and

$$I(t) = \frac{E}{R}\left(1 - e^{-Rt/L}\right). \tag{1.5.33}$$

Figure 1.5.3 illustrates Equation 1.5.33 as a function of time. Initially the current increases rapidly but the growth slows with time. Note that we could also have solved this problem by separation of variables.

Quite often, the solution is separated into two parts: the *steady-state solution* and the *transient solution*. The steady-state solution is that portion of the solution which remains as $t \to \infty$. It can equal zero. Presently it equals the constant value, E/R. The transient solution is that portion of the solution which vanishes as time increases. Here it equals $-Ee^{-Rt/L}/R$.

Although our analysis is a useful approximation to the real world, a more realistic one would include the nonlinear properties of the resistor.[9] To illustrate this, consider the case of an RL circuit without any electromotive source ($E = 0$) where the initial value for the current is I_0. Equation 1.5.29 now reads

$$L\frac{dI}{dt} + RI(1 - aI) = 0, \qquad I(0) = I_0. \tag{1.5.34}$$

Separating the variables,

$$\frac{dI}{I(aI - 1)} = \frac{dI}{I - 1/a} - \frac{dI}{I} = \frac{R}{L}dt. \tag{1.5.35}$$

[9] For the analysis of

$$L\frac{dI}{dt} + RI + KI^{\beta} = 0,$$

see Fairweather, A., and J. Ingham, 1941: Subsidence transients in circuits containing a non-linear resistor, with reference to the problem of spark-quenching. *J. IEE, Part 1*, **88**, 330–339.

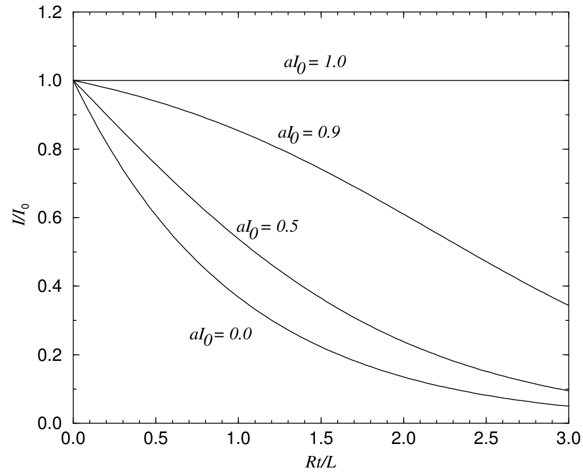


Figure 1.5.4: The variation of current I/I_0 as a function of time Rt/L with different values of aI_0.

Upon integrating and applying the initial condition, we have that

$$I = \frac{I_0 e^{-Rt/L}}{1 - aI_0 + aI_0 e^{-Rt/L}}. \qquad (\mathbf{1.5.36})$$

Figure 1.5.4 shows $I(t)$ for various values of a. As the nonlinearity reduces resistance, the decay in the current is reduced. If $aI_0 > 1$, Equation 1.5.36 predicts that the current would grow with time. The point here is that nonlinearity can have a dramatic influence on a physical system.

Consider now the electrical circuit shown in Figure 1.5.5, which contains a resistor with resistance R and a capacitor with capacitance C. Here the voltage drop across the resistor is still RI while the voltage drop across the capacitor is Q/C. Therefore, by Kirchhoff's law,

$$RI + \frac{Q}{C} = E. \qquad (\mathbf{1.5.37})$$

Equation 1.5.37 is *not* a differential equation. However, because current is the time rate of change in charge $I = dQ/dt$, our differential equation becomes

$$R\frac{dQ}{dt} + \frac{Q}{C} = E, \qquad (\mathbf{1.5.38})$$

which is the differential equation for the instantaneous charge.

Let us solve Equation 1.5.38 when the resistance and capacitance are constant but the electromotive force equals $E_0 \cos(\omega t)$. The corresponding differential equation is now

$$R\frac{dQ}{dt} + \frac{Q}{C} = E_0 \cos(\omega t). \qquad (\mathbf{1.5.39})$$

The differential equation has the integrating factor $e^{t/(RC)}$ so that it can be rewritten

$$\frac{d}{dt}\left[e^{t/(RC)} Q(t)\right] = \frac{E_0}{R} e^{t/(RC)} \cos(\omega t). \qquad (\mathbf{1.5.40})$$

Integrating Equation 1.5.40,

$$e^{t/(RC)} Q(t) = \frac{CE_0}{1 + R^2C^2\omega^2} e^{t/(RC)} \left[\cos(\omega t) + RC\omega \sin(\omega t)\right] + C_1 \qquad (\mathbf{1.5.41})$$

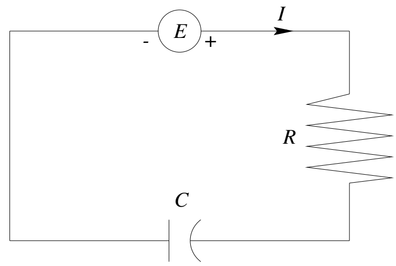


Figure 1.5.5: Schematic diagram for an electric circuit that contains a resistor of resistance R and a capacitor of capacitance C.

or

$$Q(t) = \frac{CE_0}{1 + R^2C^2\omega^2} [\cos(\omega t) + RC\omega \sin(\omega t)] + C_1 e^{-t/(RC)}. \quad \textbf{(1.5.42)}$$

If we take the initial condition as $Q(0) = 0$, then the final solution is

$$Q(t) = \frac{CE_0}{1 + R^2C^2\omega^2} \left[\cos(\omega t) - e^{-t/(RC)} + RC\omega \sin(\omega t)\right]. \quad \textbf{(1.5.43)}$$

Figure 1.5.6 illustrates Equation 1.5.43. Note how the circuit eventually supports a purely oscillatory solution (the steady-state solution) as the exponential term decays to zero (the transient solution). Indeed, the purpose of the transient solution is to allow the system to adjust from its initial condition to the final steady state. □

• Example 1.5.4: Terminal velocity

When an object passes through a fluid, the viscosity of the fluid resists the motion by exerting a force on the object proportional to its velocity. Let us find the motion of a mass m that is initially thrown upward with the speed v_0.

If we choose the coordinate system so that it increases in the vertical direction, then the equation of motion is

$$m\frac{dv}{dt} = -kv - mg \quad \textbf{(1.5.44)}$$

with $v(0) = v_0$ and $k > 0$. Rewriting Equation 1.5.44, we obtain the first-order linear differential equation

$$\frac{dv}{dt} + \frac{k}{m}v = -g. \quad \textbf{(1.5.45)}$$

Its solution in *nondimensional* form is

$$\frac{kv(t)}{mg} = -1 + \left(1 + \frac{kv_0}{mg}\right) e^{-kt/m}. \quad \textbf{(1.5.46)}$$

The displacement from its initial position is

$$\frac{k^2x(t)}{m^2g} = \frac{k^2x_0}{m^2g} - \frac{kt}{m} + \left(1 + \frac{kv_0}{mg}\right)\left(1 - e^{-kt/m}\right). \quad \textbf{(1.5.47)}$$

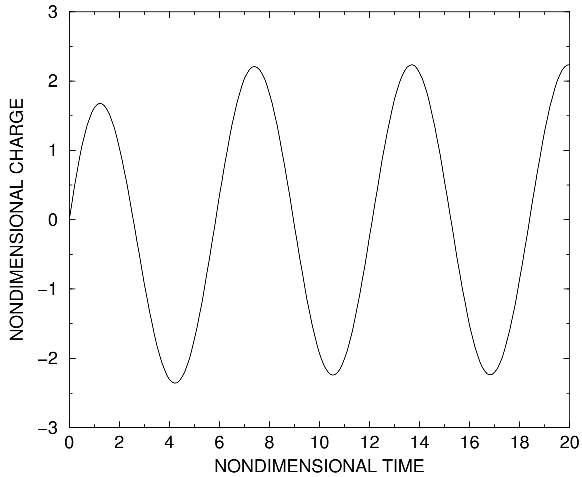


Figure 1.5.6: The temporal evolution of the nondimensional charge $(1 + R^2C^2\omega^2)Q(t)\ /(CE_0)$ in the electric circuit shown in Figure 1.5.4 as a function of nondimensional time ωt when the circuit is driven by the electromotive force $E_0 \cos(\omega t)$ and $RC\omega = 2$.

As $t \to \infty$, the velocity tends to a constant downward value, $-mg/k$, the so-called "terminal velocity," where the aerodynamic drag balances the gravitational acceleration. This is the steady-state solution.

Why have we written Equation 1.5.46 and Equation 1.5.47 in this nondimensional form? There are two reasons. First, the solution reduces to three fundamental variables, a *nondimensional* displacement $x_* = k^2x(t)/(m^2g)$, velocity $v_* = kv(t)/(mg)$, and time $t_* = kt/m$, rather than the six original parameters and variables: g, k, m, t, v, and x. Indeed, if we had substituted t_*, v_*, and x_* into Equation 1.5.45, we would have obtained the following simplified initial-value problem:

$$\frac{dv_*}{dt_*} + v_* = -1, \quad \frac{dx_*}{dt_*} = v_*, \quad v_*(0) = \frac{kv_0}{mg}, \quad x_*(0) = \frac{k^2x_0}{m^2g} \tag{1.5.48}$$

right from the start. The second advantage of the nondimensional form is the compact manner in which the results can be displayed, as Figure 1.5.7 shows.

From Equation 1.5.46 and Equation 1.5.47, the trajectory of the ball is as follows: If we define the coordinate system so that $x_0 = 0$, then the object will initially rise to the height H given by

$$\frac{k^2H}{m^2g} = \frac{kv_0}{mg} - \ln\left(1 + \frac{kv_0}{mg}\right) \tag{1.5.49}$$

at the time

$$\frac{kt_{max}}{m} = \ln\left(1 + \frac{kv_0}{mg}\right), \tag{1.5.50}$$

when $v(t_{max}) = 0$. It will then fall toward the earth. Given sufficient time $kt/m \gg 1$, it would achieve terminal velocity. □

• Example 1.5.5: The Bernoulli equation

Bernoulli's equation,

$$\frac{dy}{dx} + p(x)y = q(x)y^n, \qquad n \neq 0, 1, \tag{1.5.51}$$

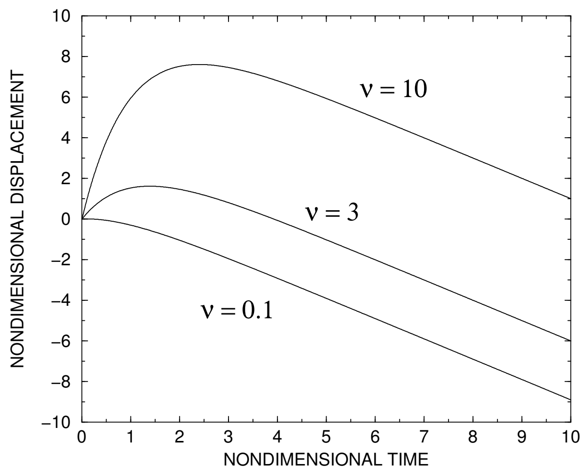


Figure 1.5.7: The nondimensional displacement $k^2x(t)/(m^2g)$ as a function of nondimensional time kt/m of an object of mass m thrown upward at the initial nondimensional speed $\nu = kv_0/(mg)$ in a fluid that retards its motion as $-kv$.

is a first-order, nonlinear differential equation. This equation can be transformed into a first-order, linear differential equation by introducing the change of variable $z = y^{1-n}$. Because

$$\frac{dz}{dx} = (1-n)y^{-n}\frac{dy}{dx}, \tag{1.5.52}$$

the transformed Bernoulli equation becomes

$$\frac{dz}{dx} + (1-n)p(x)z = (1-n)q(x). \tag{1.5.53}$$

This is now a first-order linear differential equation for z and can be solved using the methods introduced in this section. Once z is known, the solution is found by transforming back from z to y.

To illustrate this procedure, consider the nonlinear ordinary differential equation

$$x^2y\frac{dy}{dx} - xy^2 = 1, \tag{1.5.54}$$

or

$$\frac{dy}{dx} - \frac{y}{x} = \frac{y^{-1}}{x^2}. \tag{1.5.55}$$

Equation 1.5.55 is a Bernoulli equation with $p(x) = -1/x$, $q(x) = 1/x^2$, and $n = -1$. Introducing $z = y^2$, it becomes

$$\frac{dz}{dx} - \frac{2z}{x} = \frac{2}{x^2}. \tag{1.5.56}$$

This first-order linear differential equation has the integrating factor $\mu(x) = 1/x^2$ and

$$\frac{d}{dx}\left(\frac{z}{x^2}\right) = \frac{2}{x^4}. \tag{1.5.57}$$

Integration gives

$$\frac{z}{x^2} = C - \frac{2}{3x^3}. \tag{1.5.58}$$

Therefore, the general solution is

$$y^2 = z = Cx^2 - \frac{2}{3x}. \tag{1.5.59}$$

Problems

Find the solution for the following differential equations. State the interval on which the general solution is valid. Then use MATLAB to examine their behavior for a wide class of initial conditions.

1. $y' + y = e^x$
2. $y' + 2xy = x$
3. $x^2 y' + xy = 1$
4. $(2y + x^2)\,dx = x\,dy$
5. $y' - 3y/x = 2x^2$
6. $y' + 2y = 2\sin(x)$
7. $y' + 2\cos(2x)y = 0$
8. $xy' + y = \ln(x)$
9. $y' + 3y = 4, \quad y(0) = 5$
10. $y' - y = e^x/x, \quad y(e) = 0$
11. $\sin(x)y' + \cos(x)y = 1$
12. $[1 - \cos(x)]y' + 2\sin(x)y = \tan(x)$
13. $y' + [a\tan(x) + b\sec(x)]y = c\sec(x)$
14. $(xy + y - 1)\,dx + x\,dy = 0$
15. $y' + 2ay = \dfrac{x}{2} - \dfrac{\sin(2\omega x)}{4\omega}, \quad y(0) = 0.$
16. $y' + \dfrac{2k}{x^3}y = \ln\left(\dfrac{x+1}{x}\right), \ k > 0, \ y(1) = 0.$
17. Solve the following initial-value problem:

$$kxy\frac{dy}{dx} = y^2 - x, \qquad y(1) = 0.$$

Hint: Introduce the new dependent variable $p = y^2$.

18. If $x(t)$ denotes the equity capital of a company, then under certain assumptions[10] $x(t)$ is governed by

$$\frac{dx}{dt} = (1 - N)rx + S,$$

where N is the dividend payout ratio, r is the rate of return of equity, and S is the rate of net new stock financing. If the initial value of $x(t)$ is $x(0)$, find $x(t)$.

19. The assimilation[11] of a drug into a body can be modeled by the chemical reaction A $\stackrel{k_1}{\to}$ B $\stackrel{k_2}{\to}$ C, which is governed by the chemical kinetics equations

$$\frac{d[\mathrm{A}]}{dt} = -k_1[\mathrm{A}], \qquad \frac{d[\mathrm{B}]}{dt} = k_1[\mathrm{A}] - k_2[\mathrm{B}], \qquad \frac{d[\mathrm{C}]}{dt} = k_2[\mathrm{B}],$$

---

[10] See Lebowitz, J. L., C. O. Lee, and P. B. Linhart, 1976: Some effects of inflation on a firm with original cost depreciation. *Bell J. Economics*, **7**, 463–477.

[11] See Calder, G. V., 1974: The time evolution of drugs in the body: An application of the principle of chemical kinetics. *J. Chem. Educ.*, **51**, 19–22.

where [A] denotes the concentration of the drug in the gastrointestinal tract or in the site of injection, [B] is the concentration of the drug in the body, and [C] is either the amount of drug eliminated by various metabolic functions or the amount of the drug utilized by various action sites in the body. If $[A]_0$ denotes the initial concentration of A, find [A], [B], and [C] as a function of time t.

20. Find the current in an RL circuit when the electromotive source equals $E_0 \cos^2(\omega t)$. Initially the circuit is dead.

Find the general solution for the following Bernoulli equations:

21. $\dfrac{dy}{dx} + \dfrac{y}{x} = -y^2$

22. $x^2 \dfrac{dy}{dx} = xy + y^2$

23. $\dfrac{dy}{dx} - \dfrac{4y}{x} = x\sqrt{y}$

24. $\dfrac{dy}{dx} + \dfrac{y}{x} = -xy^2$

25. $2xy\dfrac{dy}{dx} - y^2 + x = 0$

26. $x\dfrac{dy}{dx} + y = \frac{1}{2}xy^3$

1.6 GRAPHICAL SOLUTIONS

In spite of the many techniques developed for their solution, many ordinary differential equations cannot be solved analytically. In the next two sections, we highlight two alternative methods when analytical methods fail. Graphical methods seek to understand the nature of the solution by examining the differential equations at various points and infer the complete solution from these results. In the last section, we highlight the numerical techniques that are now commonly used to solve ordinary differential equations on the computer.

- *Direction fields*

One of the simplest numerical methods for solving first-order ordinary differential equations follows from the fundamental concept that the derivative gives the *slope* of a straight line that is tangent to a curve at a given point.

Consider the first-order differential equation

$$y' = f(x, y), \tag{1.6.1}$$

which has the initial value $y(x_0) = y_0$. For any (x, y) it is possible to draw a short line segment whose slope equals $f(x, y)$. This graphical representation is known as the *direction field* or *slope field* of Equation 1.6.1. Starting with the initial point (x_0, y_0), we can then construct the *solution curve* by extending the initial line segment in such a manner that the tangent of the solution curve parallels the direction field at each point through which the curve passes.

Before the days of computers, it was common to first draw lines of constant slope (*isoclines*) or $f(x, y) = c$. Because along any isocline all of the line segments had the same slope, considerable computational savings were realized. Today, computer software exists that performs these graphical computations with great speed.

To illustrate this technique, consider the ordinary differential equation

$$\frac{dx}{dt} = x - t^2. \tag{1.6.2}$$

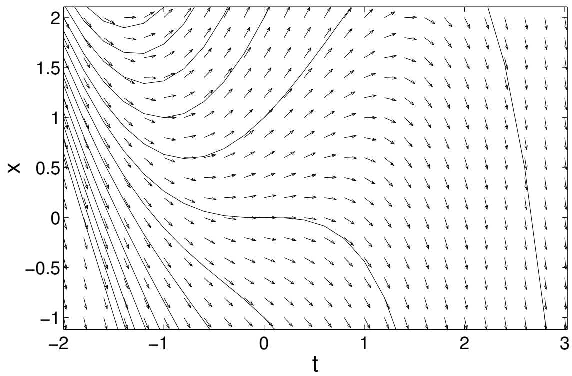


Figure 1.6.1: The direction field for Equation 1.6.2. The solid lines are plots of the solution with various initial conditions.

Its exact solution is

$$x(t) = Ce^t + t^2 + 2t + 2, \tag{1.6.3}$$

where C is an arbitrary constant. Using the MATLAB script

```
clear
% create grid points in t and x
[t,x] = meshgrid(-2:0.2:3,-1:0.2:2);
% load in the slope
slope = x - t.*t;
% find the length of the vector (1,slope)
length = sqrt(1 + slope .* slope);
% create and plot the vector arrows
quiver(t,x,1./length,slope./length,0.5)
axis equal tight
hold on
% plot the exact solution for various initial conditions
tt = [-2:0.2:3];
for cval = -10:1:10
  x_exact = cval * exp(tt) + tt.*tt + 2*tt + 2;
  plot(tt,x_exact)
  xlabel('t','Fontsize',20)
  ylabel('x','Fontsize',20)
end
```

we show in Figure 1.6.1 the directional field associated with Equation 1.6.2 along with some of the particular solutions. Clearly the vectors are parallel to the various particular solutions. Therefore, without knowing the solution, we could choose an arbitrary initial condition and sketch its behavior at subsequent times. The same holds true for nonlinear equations.

• *Rest points and autonomous equations*

In the case of autonomous differential equations (equations where the independent variable does not explicitly appear in the equation), considerable information can be gleaned from a graphical analysis of the equation.

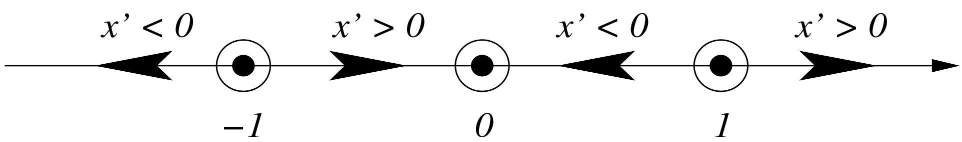


Figure 1.6.2: The phase line diagram for the ordinary differential equation, Equation 1.6.4.

Consider the nonlinear ordinary differential equation

$$x' = \frac{dx}{dt} = x(x^2 - 1). \qquad \textbf{(1.6.4)}$$

The time derivative x' vanishes at $x = -1, 0, 1$. Consequently, if $x(0) = 0$, $x(t)$ will remain zero forever. Similarly, if $x(0) = 1$ or $x(0) = -1$, then $x(t)$ will equal 1 or -1 for all time. For this reason, values of x for which the derivative x' is zero are called *rest points*, *equilibrium points*, or *critical points* of the differential equation.

The behavior of solutions near rest points is often of considerable interest. For example, what happens to the solution when x is near one of the rest points $x = -1, 0, 1$?

Consider the point $x = 0$. For x slightly greater than zero, $x' < 0$. For x slightly less than 0, $x' > 0$. Therefore, for any initial value of x near $x = 0$, x will tend to zero. In this case, the point $x = 0$ is an asymptotically *stable critical point* because whenever x is perturbed away from the critical point, it tends to return there again.

Turning to the point $x = 1$, for x slightly greater than 1, $x' > 0$; for x slightly less than 1, $x' < 0$. Because any x near $x = 1$, but not equal to 1, will move away from $x = 1$, the point $x = 1$ is called an *unstable critical point*. A similar analysis applies at the point $x = -1$. This procedure of determining the behavior of an ordinary differential equation near its critical points is called a *graphical stability analysis*.

• *Phase line*

A graphical representation of the results of our graphical stability analysis is the *phase line*. On a phase line, the equilibrium points are denoted by circles. See Figure 1.6.2. Also on the phase line we identify the sign of x' for all values of x. From the sign of x', we then indicate whether x is increasing or deceasing by an appropriate arrow. If the arrow points toward the right, x is increasing; toward the left x decreases. Then, by knowing the *sign* of the derivative for all values of x, together with the starting value of x, we can determine what happens as $t \to \infty$. Any solution that is approached asymptotically as $t \to \infty$ is called a *steady-state output*. In our present example, $x = 0$ is a steady-state output.

Problems

In previous sections, you used various techniques to solve first-order ordinary differential equations. Now check your work by using MATLAB to draw the direction field and plot your analytic solution for the following problems taken from previous sections:

1. Section 1.2, Problem 5
2. Section 1.3, Problem 1
3. Section 1.4, Problem 5
4. Section 1.5, Problem 3

For the following autonomous ordinary differential equations, draw the phase line. Then classify each equilibrium solution as either stable or unstable.

5. $x' = \alpha x(1-x)(x-\frac{1}{2})$

6. $x' = (x^2-1)(x^2-4)$

7. $x' = -4x - x^3$

8. $x' = 4x - x^3$

1.7 NUMERICAL METHODS

By now you have seen most of the exact methods for finding solutions to first-order ordinary differential equations. The methods have also given you a view of the general behavior and properties of solutions to differential equations. However, it must be admitted that in many instances exact solutions cannot be found and we must resort to numerical solutions.

In this section we present the two most commonly used methods for solving differential equations: Euler and Runge-Kutta methods. There are many more methods and the interested student is referred to one of countless numerical methods books. A straightforward extension of these techniques can be applied to systems of first-order and higher-order differential equations.

• *Euler and modified Euler methods*

Consider the following first-order differential equation and initial condition:

$$\frac{dy}{dx} = f(x,y), \qquad y(x_0) = y_0. \tag{1.7.1}$$

Euler's method is based on a Taylor series expansion of the solution about x_0 or

$$y(x_0+h) = y(x_0) + hy'(x_0) + \tfrac{1}{2}y''(\xi)h^2, \qquad x_0 < \xi < x_0 + h, \tag{1.7.2}$$

where h is the step size. Euler's method consists of taking a sufficiently small h so that only the first two terms of this Taylor expansion are significant.

Let us now replace $y'(x_0)$ by $f(x_0, y_0)$. Using subscript notation, we have that

$$y_{i+1} = y_i + hf(x_i, y_i) + O(h^2). \tag{1.7.3}$$

Equation 1.7.3 states that if we know the values of y_i and $f(x_i, y_i)$ at the position x_i, then the solution at x_{i+1} can be obtained with an error[12] $O(h^2)$.

The trouble with Euler's method is its lack of accuracy, often requiring an extremely small time step. How might we improve this method with little additional effort?

One possible method would retain the first three terms of the Taylor expansion rather than the first two. This scheme, known as the *modified Euler method*, is

$$y_{i+1} = y_i + hy'(x_i) + \tfrac{1}{2}h^2 y_i'' + O(h^3). \tag{1.7.4}$$

This is clearly more accurate than Equation 1.7.3.

---

[12] The symbol O is a mathematical notation indicating relative magnitude of terms, namely that $f(\epsilon) = O(\epsilon^n)$ provided $\lim_{\epsilon\to 0}|f(\epsilon)/\epsilon^n| < \infty$. For example, as $\epsilon \to 0$, $\sin(\epsilon) = O(\epsilon)$, $\sin(\epsilon^2) = O(\epsilon^2)$, and $\cos(\epsilon) = O(1)$.

An obvious question is how do we evaluate y_i'', because we do not have any information on its value. Using the forward derivative approximation, we find that

$$y_i'' = \frac{y_{i+1}' - y_i'}{h}. \tag{1.7.5}$$

Substituting Equation 1.7.5 into Equation 1.7.4 and simplifying

$$y_{i+1} = y_i + \frac{h}{2}\left(y_i' + y_{i+1}'\right) + O(h^3). \tag{1.7.6}$$

Using the differential equation,

$$y_{i+1} = y_i + \frac{h}{2}\left[f(x_i, y_i) + f(x_{i+1}, y_{i+1})\right] + O(h^3). \tag{1.7.7}$$

Although $f(x_i, y_i)$ at (x_i, y_i) are easily calculated, how do we compute $f(x_{i+1}, y_{i+1})$ at (x_{i+1}, y_{i+1})? For this we compute a first guess via the Euler method, Equation 1.7.3; Equation 1.7.7 then provides a refinement on the value of y_{i+1}.

In summary then, the simple Euler scheme is

$$y_{i+1} = y_i + k_1 + O(h^2), \quad k_1 = hf(x_i, y_i), \tag{1.7.8}$$

while the modified Euler method is

$$y_{i+1} = y_i + \tfrac{1}{2}(k_1 + k_2) + O(h^3),\ k_1 = hf(x_i, y_i),\ k_2 = hf(x_i + h, y_i + k_1). \tag{1.7.9}$$

• **Example 1.7.1**

Let us illustrate Euler's method by numerically solving

$$x' = x + t, \qquad x(0) = 1. \tag{1.7.10}$$

A quick check shows that Equation 1.7.10 has the exact solution $x_{\text{exact}}(t) = 2e^t - t - 1$. Using the MATLAB script

```
clear
for i = 1:3
% set up time step increment and number of time steps
  h = 1/10^i; n = 10/h;
% set up initial conditions
  t=zeros(n+1,1); t(1) = 0;
  x_euler=zeros(n+1,1); x_euler(1) = 1;
  x_modified=zeros(n+1,1); x_modified(1) = 1;
  x_exact=zeros(n+1,1); x_exact(1) = 1;
% set up difference arrays for plotting purposes
  diff1 = zeros(n,1); diff2 = zeros(n,1); tplot = zeros(n,1);
% define right side of differential equation, Equation 1.7.10
  f = inline('xx+tt','tt','xx');
  for k = 1:n
    t(k+1) = t(k) + h;
% compute exact solution
```

```
    x_exact(k+1) = 2*exp(t(k+1)) - t(k+1) - 1;
% compute solution via Euler's method
    k1 = h * f(t(k),x_euler(k));
    x_euler(k+1) = x_euler(k) + k1;
    tplot(k) = t(k+1);
    diff1(k) = x_euler(k+1) - x_exact(k+1);
    diff1(k) = abs(diff1(k) / x_exact(k+1));
% compute solution via modified Euler method
    k1 = h * f(t(k),x_modified(k));
    k2 = h * f(t(k+1),x_modified(k)+k1);
    x_modified(k+1) = x_modified(k) + 0.5*(k1+k2);
    diff2(k) = x_modified(k+1) - x_exact(k+1);
    diff2(k) = abs(diff2(k) / x_exact(k+1));
  end
% plot relative errors
  semilogy(tplot,diff1,'-',tplot,diff2,':')
  hold on
  xlabel('TIME','Fontsize',20)
  ylabel('|RELATIVE ERROR|','Fontsize',20)
  legend('Euler method','modified Euler method')
  legend boxoff;
  num1 = 0.2*n; num2 = 0.8*n;
  text(3,diff1(num1),['h = ',num2str(h)],'Fontsize',15,...
      'HorizontalAlignment','right',...
      'VerticalAlignment','bottom')
  text(9,diff2(num2),['h = ',num2str(h)],'Fontsize',15,...
      'HorizontalAlignment','right',...
      'VerticalAlignment','bottom')
end
```

Both the Euler and modified Euler methods have been used to numerically integrate Equation 1.7.10 and the absolute value of the relative error is plotted in Figure 1.7.1 as a function of time for various time steps. In general, the error grows with time. The decrease of error with smaller time steps, as predicted in our analysis, is quite apparent. Furthermore, the superiority of the modified Euler method over the original Euler method is clearly seen. □

• *Runge-Kutta method*

As we have just shown, the accuracy of numerical solutions of ordinary differential equations can be improved by adding more terms to the Taylor expansion. The Runge-Kutta method[13] builds upon this idea, just as the modified Euler method did.

Let us assume that the numerical solution can be approximated by

$$y_{i+1} = y_i + ak_1 + bk_2, \tag{1.7.11}$$

[13] Runge, C., 1895: Ueber die numerische Auflösung von Differentialgleichungen. *Math. Ann.*, **46**, 167–178; Kutta, W., 1901: Beitrag zur Näherungsweisen Integration totaler Differentialgleichungen. *Zeit. Math. Phys.*, **46**, 435–453. For a historical review, see Butcher, J. C., 1996: A history of Runge-Kutta methods. *Appl. Numer. Math.*, **20**, 247–260 and Butcher, J. C., and G. Wanner, 1996: Runge-Kutta methods: Some historical notes. *Appl. Numer. Math.*, **22**, 113–151.

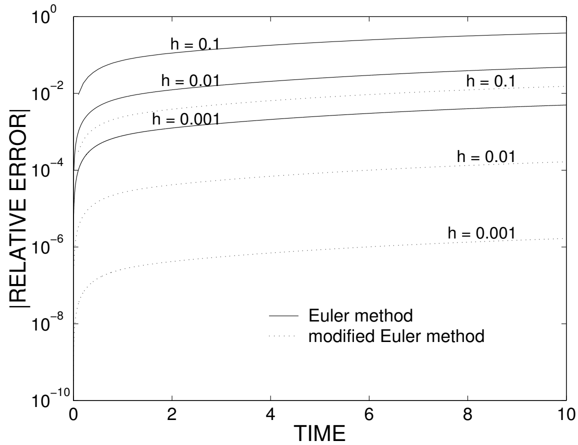


Figure 1.7.1: The relative error $[x(t) - x_{\text{exact}}(t)]/x_{\text{exact}}(t)$ of the numerical solution of Equation 1.7.10 using Euler's method (the solid line) and modified Euler's method (the dotted line) with different time steps h.

where

$$k_1 = hf(x_i, y_i) \qquad \text{and} \qquad k_2 = hf(x_i + A_1 h, y_i + B_1 k_1). \tag{1.7.12}$$

Here a, b, A_1, and B_1 are four unknowns. Equation 1.7.11 was suggested by the modified Euler method that we just presented. In that case, the truncated Taylor series had an error of $O(h^3)$. We anticipate such an error in the present case.

Because the Taylor series expansion of $f(x+h, y+k)$ about (x, y) is

$$\begin{aligned} f(x+h, y+k) &= f(x,y) + (hf_x + kf_y) + \tfrac{1}{2}\left(h^2 f_{xx} + 2hkf_{xy} + k^2 f_{yy}\right) \\ &\quad + \tfrac{1}{6}\left(h^3 f_{xxx} + 3h^2 k f_{xxy} + 3hk^2 f_{xyy} + k^3 f_{yyy}\right) + \cdots, \end{aligned} \tag{1.7.13}$$

k_2 can be rewritten

$$k_2 = hf[x_i + A_1 h, y_i + Bhf(x_i, y_i)] \tag{1.7.14}$$

$$= h\left[f(x_i, y_i) + (A_1 h f_x + B_1 h f f_y)\right] \tag{1.7.15}$$

$$= hf + A_1 h^2 f_x + B_1 h^2 f f_y, \tag{1.7.16}$$

where we have retained only terms up to $O(h^2)$ and neglected all higher-order terms. Finally, substituting Equation 1.7.16 into Equation 1.7.11 gives

$$y_{i+1} = y_i + (a+b)hf + (A_1 b f_x + B_1 b f f_y)h^2. \tag{1.7.17}$$

This equation corresponds to the second-order Taylor expansion:

$$y_{i+1} = y_i + hy_i' + \tfrac{1}{2}h^2 y_i''. \tag{1.7.18}$$

Therefore, if we wish to solve the differential equation $y' = f(x, y)$, then

$$y'' = f_x + f_y y' = f_x + f f_y. \tag{1.7.19}$$

Substituting Equation 1.7.19 into Equation 1.7.18, we have that

$$y_{i+1} = y_i + hf + \tfrac{1}{2}h^2(f_x + f f_y). \tag{1.7.20}$$

Although Carl David Tolmé Runge (1856–1927) began his studies in Munich, his friendship with Max Planck led him to Berlin and pure mathematics with Kronecker and Weierstrass. It was his professorship at Hanover beginning in 1886 and subsequent work in spectroscopy that led him to his celebrated paper on the numerical integration of ordinary differential equations. Runge's final years were spent in Göttingen as a professor in applied mathematics. (Portrait taken with permission from Reid, C., 1976: *Courant in Göttingen and New York: The Story of an Improbable Mathematician.* Springer-Verlag, 314 pp. ©1976, by Springer-Verlag New York Inc.)

A direct comparison of Equation 1.7.17 and Equation 1.7.20 yields

$$a + b = 1, \qquad A_1 b = \tfrac{1}{2}, \quad \text{and} \quad B_1 b = \tfrac{1}{2}. \tag{1.7.21}$$

These three equations have four unknowns. If we choose $a = \frac{1}{2}$, we immediately calculate $b = \frac{1}{2}$ and $A_1 = B_1 = 1$. Hence the second-order Runge-Kutta scheme is

$$y_{i+1} = y_i + \tfrac{1}{2}(k_1 + k_2), \tag{1.7.22}$$

where $k_1 = hf(x_i, y_i)$ and $k_2 = hf(x_i + h, y_i + k_1)$. Thus, the second-order Runge-Kutta scheme is identical to the modified Euler method.

Although the derivation of the second-order Runge-Kutta scheme yields the modified Euler scheme, it does provide a framework for computing higher-order and more accurate schemes. A particularly popular one is the fourth-order Runge-Kutta scheme

$$y_{i+1} = y_i + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \tag{1.7.23}$$

where

$$k_1 = hf(x_i, y_i), \tag{1.7.24}$$

$$k_2 = hf(x_i + \tfrac{1}{2}h, y_i + \tfrac{1}{2}k_1), \tag{1.7.25}$$

$$k_3 = hf(x_i + \tfrac{1}{2}h, y_i + \tfrac{1}{2}k_2), \tag{1.7.26}$$

and

$$k_4 = hf(x_i + h, y_i + k_3). \tag{1.7.27}$$

• **Example 1.7.2**

Let us illustrate the fourth-order Runge-Kutta by redoing the previous example using the MATLAB script

```
clear
% test out different time steps
for i = 1:4
% set up time step increment and number of time steps
  if i==1 h = 0.50; end; if i==2 h = 0.10; end;
  if i==3 h = 0.05; end; if i==4 h = 0.01; end;
  n = 10/h;
% set up initial conditions
  t=zeros(n+1,1); t(1) = 0;
  x_rk=zeros(n+1,1); x_rk(1) = 1;
  x_exact=zeros(n+1,1); x_exact(1) = 1;
% set up difference arrays for plotting purposes
  diff = zeros(n,1); tplot = zeros(n,1);
% define right side of differential equation
  f = inline('xx+tt','tt','xx');
  for k = 1:n
    x_local = x_rk(k); t_local = t(k);
    k1 = h * f(t_local,x_local);
    k2 = h * f(t_local + h/2,x_local + k1/2);
    k3 = h * f(t_local + h/2,x_local + k2/2);
    k4 = h * f(t_local + h,x_local + k3);
    t(k+1) = t_local + h;
    x_rk(k+1) = x_local + (k1+2*k2+2*k3+k4) / 6;
    x_exact(k+1) = 2*exp(t(k+1)) - t(k+1) - 1;
    tplot(k) = t(k);
    diff(k) = x_rk(k+1) - x_exact(k+1);
    diff(k) = abs(diff(k) / x_exact(k+1));
  end
% plot relative errors
  semilogy(tplot,diff,'-')
  hold on
  xlabel('TIME','Fontsize',20)
  ylabel('|RELATIVE ERROR|','Fontsize',20)
  num1 = 2*i; num2 = 0.2*n;
  text(num1,diff(num2),['h = ',num2str(h)],'Fontsize',15,...
      'HorizontalAlignment','right',...
      'VerticalAlignment','bottom')
end
```

The error growth with time is shown in Figure 1.7.2. Although this script could be used for any first-order ordinary differential equation, the people at MATLAB have an alternative called `ode45`, which combines a fourth-order and a fifth-order method that are similar to our fourth-order Runge-Kutta method. Their scheme is more efficient because it varies the

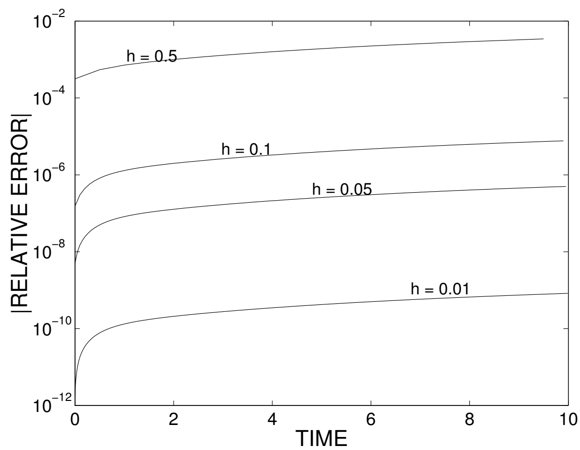


Figure 1.7.2: Same as Figure 1.7.1 except that we have used the fourth-order Runge-Kutta method.

step size, choosing a new time step at each step in an attempt to achieve a given desired accuracy.

□

• *Adams-Bashforth method*

All of the methods presented so far (Euler, modified Euler, Runge-Kutta) are single point methods; the solution at $i+1$ depends solely on a single point i. A popular alternative to these schemes are multistep methods that compute y_{i+1} by reusing previously obtained values of y_n where $n < i$.

We begin our derivation of a multistep method by rewriting Equation 1.7.1 as

$$dy = f(x, y)\,dx. \tag{1.7.28}$$

Integrating both sides of Equation 1.7.28, we obtain

$$y(x_{i+1}) - y(x_i) = \int_{x_i}^{x_{i+1}} dy = \int_{x_i}^{x_{i+1}} f(x, y)\,dx. \tag{1.7.29}$$

The Adams-Bashforth method[14] replaces the integrand in Equation 1.7.29 with an approximation derived from Newton's backward difference formula:

$$f(x, y) \approx f_i + \xi\nabla f_i + \tfrac{1}{2}\xi(\xi+1)\nabla^2 f_i + \tfrac{1}{6}\xi(\xi+1)(\xi+2)\nabla^3 f_i, \tag{1.7.30}$$

where $\xi = (x - x_i)/h$ or $x = x_i + h\xi$,

$$\nabla f_i = f(x_i, y_i) - f(x_{i-1}, y_{i-1}), \tag{1.7.31}$$

$$\nabla^2 f_i = f(x_i, y_i) - 2f(x_{i-1}, y_{i-1}) + f(x_{i-2}, y_{i-2}), \tag{1.7.32}$$

---

[14] Bashforth, F., and J. C. Adams, 1883: *An Attempt to Test the Theories of Capillary Action by Comparing the Theoretical and Measured Forms of Drops of Fluid. With an Explanation of the Method of Integration Employed in Constructing the Tables Which Give the Theoretical Forms of Such Drops.* Cambridge University Press, 139 pp.

and

$$\nabla^3 f_i = f(x_i, y_i) - 3f(x_{i-1}, y_{i-1}) + 3f(x_{i-2}, y_{i-2}) - f(x_{i-3}, y_{i-3}). \quad \textbf{(1.7.33)}$$

Substituting Equation 1.7.30 into Equation 1.7.29 and carrying out the integration, we find that

$$y(x_{i+1}) = y(x_i) + \frac{h}{24}\left[55f(x_i, y_i) - 59f(x_{i-1}, y_{i-1}) + 37f(x_{i-2}, y_{i-2}) - 9f(x_{i-3}, y_{i-3})\right]. \quad \textbf{(1.7.34)}$$

Thus, the Adams-Bashforth method is an explicit finite difference formula that has a global error of $O(h^4)$. Additional computational savings can be realized if the old values of the slope are stored and used later. A disadvantage is that some alternative scheme (usually Runge-Kutta) must provide the first three starting values.

• Example 1.7.3

The flight of projectiles provides a classic application of first-order differential equations. If the projectile has a mass m and its motion is opposed by the drag $mgkv^2$, where g is the acceleration due to gravity and k is the quadratic drag coefficients, Newton's law of motion gives

$$\frac{dv}{dt} = -g\sin(\theta) - gkv^2, \quad \textbf{(1.7.35)}$$

where θ is the slope of the trajectory to the horizon. From kinematics,

$$\frac{dx}{dt} = v\cos(\theta), \quad \frac{dy}{dt} = v\sin(\theta), \quad \frac{d\theta}{dt} = -\frac{g\cos(\theta)}{v}. \quad \textbf{(1.7.36)}$$

An interesting aspect of this problem is the presence of a *system* of ordinary differential equations.

Although we can obtain an exact solution to this problem,[15] let us illustrate the Adams-Bashforth method to compute the solution to Equation 1.7.35 and Equation 1.7.36. We begin by computing the first three time steps using the Runge-Kutta method. Note that we first compute the k_1 for *all* of the dependent variables before we start computing the values of k_2. Similar considerations hold for k_3 and k_4.

```
clear
a = 0; b = 7.85; N = 100; g = 9.81; c = 0.000548;
h = (b-a)/N; t = (a:h:b+h);
% set initial conditions
v(1) = 44.69; theta(1) = pi/3; x(1) = 0; y(1) = 0;
for i = 1:3
  angle = theta(i); vv = v(i);
  k1_vel = -g*sin(angle) - g*c*vv*vv;
  k1_angle = -g*cos(angle) / vv;
  k1_x = vv * cos(angle);
  k1_y = vv * sin(angle);
  angle = theta(i)+h*k1_angle/2; vv = v(i)+h*k1_vel/2;
  k2_vel = -g*sin(angle) - g*c*vv*vv;
```

---

[15] Tan, A., C. H. Frick, and O. Castillo, 1987: The fly ball trajectory: An older approach revisited. *Am. J. Phys.*, **55**, 37–40; Chudinov, P. S., 2001: The motion of a point mass in a medium with a square law of drag. *J. Appl. Math. Mech.*, **65**, 421–426.

```
  k2_angle = -g*cos(angle) / vv;
  k2_x = vv * cos(angle);
  k2_y = vv * sin(angle);
  angle = theta(i)+h*k2_angle/2; vv = v(i)+h*k2_vel/2;
  k3_vel = -g*sin(angle) - g*c*vv*vv;
  k3_angle = -g*cos(angle) / vv;
  k3_x = vv * cos(angle);
  k3_y = vv * sin(angle);
  angle = theta(i)+h*k3_angle; vv = v(i)+h*k3_vel;
  k4_vel = -g*sin(angle) - g*c*vv*vv;
  k4_angle = -g*cos(angle) / vv;
  k4_x = vv * cos(angle);
  k4_y = vv * sin(angle);
  v(i+1) = v(i) + h*(k1_vel+2*k2_vel+2*k3_vel+k4_vel)/6;
  x(i+1) = x(i) + h*(k1_x+2*k2_x+2*k3_x+k4_x)/6;
  y(i+1) = y(i) + h*(k1_y+2*k2_y+2*k3_y+k4_y)/6;
  theta(i+1) = theta(i) + h*(k1_angle+2*k2_angle ...
             +2*k3_angle+k4_angle)/6;
end
```

Having computed the first three values of each of the dependent variables, we turn to the Adams-Bashforth method to compute the remaining portion of the numerical solution:

```
for i = 4:N
  angle = theta(i); vv = v(i);
  k1_vel = -g*sin(angle) - g*c*vv*vv;
  k1_angle = -g*cos(angle) / vv;
  k1_x = vv * cos(angle);
  k1_y = vv * sin(angle);
  angle = theta(i-1); vv = v(i-1);
  k2_vel = -g*sin(angle) - g*c*vv*vv;
  k2_angle = -g*cos(angle) / vv;
  k2_x = vv * cos(angle);
  k2_y = vv * sin(angle);
  angle = theta(i-2); vv = v(i-2);
  k3_vel = -g*sin(angle) - g*c*vv*vv;
  k3_angle = -g*cos(angle) / vv;
  k3_x = vv * cos(angle);
  k3_y = vv * sin(angle);
  angle = theta(i-3); vv = v(i-3);
  k4_vel = -g*sin(angle) - g*c*vv*vv;
  k4_angle = -g*cos(angle) / vv;
  k4_x = vv * cos(angle);
  k4_y = vv * sin(angle);
% Use Equation 1.7.35 and Equation 1.7.36 for v, x, y and θ
  v(i+1) = v(i) + h*(55*k1_vel-59*k2_vel+37*k3_vel-9*k4_vel)/24;
  x(i+1) = x(i) + h*(55*k1_x-59*k2_x+37*k3_x-9*k4_x)/24;
  y(i+1) = y(i) + h*(55*k1_y-59*k2_y+37*k3_y-9*k4_y)/24;
  theta(i+1) = theta(i) + h*(55*k1_angle-59*k2_angle ...
             +37*k3_angle-9*k4_angle)/24;
end
```

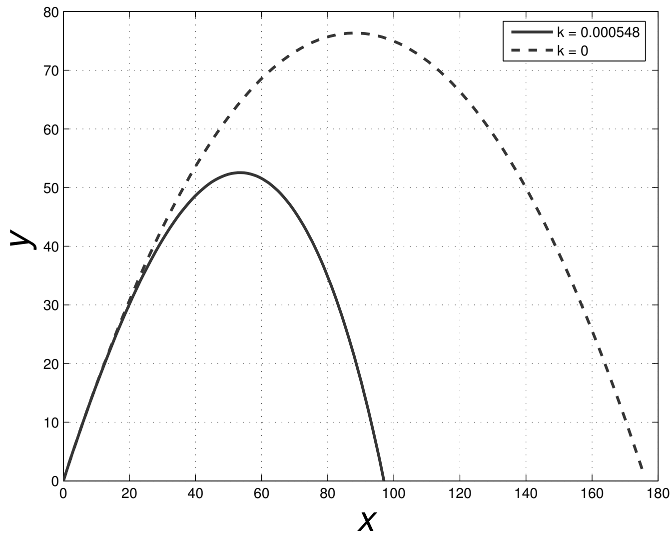


Figure 1.7.3: The trajectory of a projectile with and without air resistance when it is initial fired with a muzzle velocity of 44.69 m/s and an elevation of $\theta = 60^\circ$. All units are in the MKS system.

Figure 1.7.3 illustrates this numerical solution when $k = 0$ and $k = 0.000548$ s^2/m^2 and the shell is fired with the initial velocity $v(0) = 44.69$ m/s and elevation $\theta(0) = \pi/3$ with $x(0) = y(0) = 0$.

Problems

Using Euler's, Runge-Kutta, or the Adams-Bashforth method for various values of $h = 10^{-n}$, find the numerical solution for the following initial-value problems. Check your answer by finding the exact solution.

1. $x' = x - t, \quad x(0) = 2$

2. $x' = tx, \quad x(0) = 1$

3. $x' = x^2/(t+1), \quad x(0) = 1$

4. $x' = x + e^{-t}, \quad x(1) = 0$

5. Consider the integro-differential equation

$$\frac{dx}{dt} + \int_0^t x(\tau)\,d\tau + B\,\text{sgn}(x)|x|^\beta = 1, \qquad B, \beta \geq 0,$$

where the signum function is defined by Equation 11.2.11. This equation describes the (nondimensional) current,[16] $x(t)$, within an electrical circuit that contains a capacitor, inductor, and nonlinear resistor. Assuming that the circuit is initially dead, x(0) = 0, write a MATLAB script that uses Euler's method to compute $x(t)$. Use a simple Riemann sum to approximate the integral. See Figure 1.7.4. Examine the solution for various values of B and β as well as time step Δt.

[16] Monahan, T. F., 1960: Calculation of the current in non-linear surge-current-generation circuits. *Proc. IEE, Part C*, **107**, 288–291.

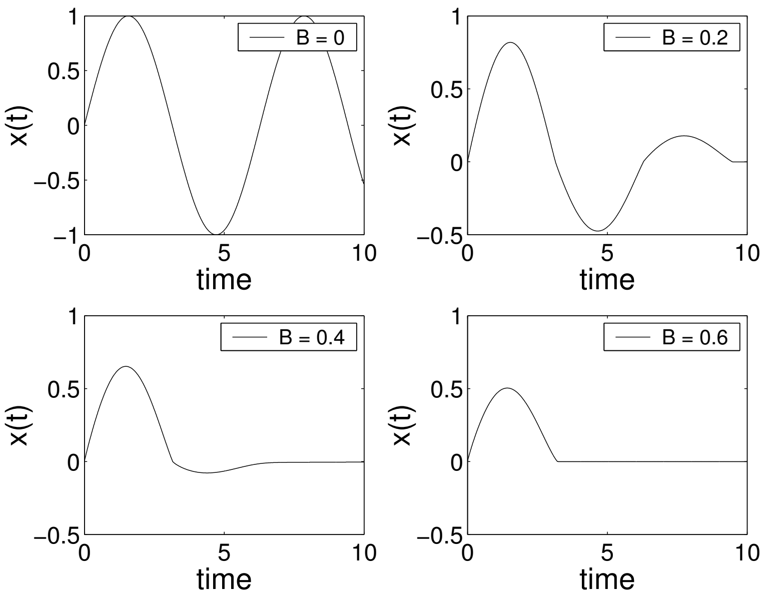


Figure 1.7.4: The numerical solution of the equation describing an electrical circuit with a nonlinear resistor. Here $\beta = 0.2$ and $\Delta t = 0.01$.

Further Readings

Boyce, W. E., and R. C. DiPrima, 2004: *Elementary Differential Equations and Boundary Value Problems.* Wiley, 800 pp. Classic textbook.

Ince, E. L., 1956: *Ordinary Differential Equations.* Dover, 558 pp. The source book on ordinary differential equations.

Zill, D. G., and M. R. Cullen, 2008: *Differential Equations with Boundary-Value Problems.* Brooks Cole, 640 pp. Nice undergraduate textbook.

Chapter 2
Higher-Order Ordinary Differential Equations

Although first-order ordinary differential equations exhibit most of the properties of differential equations, higher-order ordinary differential equations are more ubiquitous in the sciences and engineering. This chapter is devoted to the most commonly employed techniques for their solution.

A *linear* nth-order ordinary differential equation is a differential equation of the form

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = f(x). \qquad \textbf{(2.0.1)}$$

If $f(x) = 0$, then Equation 2.0.1 is said to be *homogeneous*; otherwise, it is *nonhomogeneous*. A linear differential equation is *normal* on an interval I if its coefficients and $f(x)$ are continuous, and the value of $a_n(x)$ is never zero on I.

Solutions to Equation 2.0.1 generally must satisfy not only the differential equations but also certain specified conditions at one or more points. *Initial-value problems* are problems where *all* of the conditions are specified at a single point $x = a$ and have the form: $y(a) = b_0$, $y'(a) = b_1$, $y''(a) = b_2$, ..., $y^{(n-1)}(a) = b_{n-1}$, where b_0, b_1, b_2, ..., b_{n-1} are arbitrary constants. A quick check shows that if Equation 2.0.1 is homogeneous and normal on an interval I and *all* of the initial conditions equal zero at the point $x = a$ that lies in I, then $y(x) \equiv 0$ on I. This follows because $y = 0$ is a solution of Equation 2.0.1 and satisfies the initial conditions.

At this point a natural question would be whether the solution exists for this initial-value problem and, if so, whether it is unique. From a detailed study of this question,[1] we have the following useful theorem.

Theorem: Existence and Uniqueness

Suppose that the differential equation, Equation 2.0.1, is normal on the open interval I containing the point $x = a$. Then, given n numbers b_0, b_1, ..., b_{n-1}, the nth-order linear equation, Equation 2.0.1, has a unique solution on the entire interval I that satisfies the n initial conditions $y(a) = b_0, y'(a) = b_1, \ldots, y^{(n-1)}(a) = b_{n-1}$. □

• **Example 2.0.1**

The solution $y(x) = \frac{4}{3}e^x - \frac{1}{3}e^{-2x}$ to the ordinary differential equation $y''' + 2y'' - y' - 2y = 0$ satisfies the initial conditions $y(0) = 1$, $y'(0) = 2$, and $y''(0) = 0$ at $x = 0$. Our theorem guarantees us that this is the *only* solution with *these* initial values. □

Another class of problems, commonly called (two-point) *boundary-value problems*, occurs when conditions are specified at two *different* points $x = a$ and $x = b$ with $b > a$. An important example, in the case of second-order ordinary differential equations, is the Sturm-Liouville problem where the boundary conditions are $\alpha_1 y(a) + \beta_1 y'(a) = 0$ at $x = a$ and $\alpha_2 y(b) + \beta_2 y'(b) = 0$ at $x = b$. The Sturm-Liouville problem is treated in Chapter 6.

Having introduced some of the terms associated with higher-order ordinary linear differential equations, how do we solve them? One way is to recognize that these equations are really a set of linear, first-order ordinary differential equations. For example, the linear second-order linear differential equation

$$y'' - 3y' + 2y = 3x \tag{2.0.2}$$

can be rewritten as the following system of first-order ordinary differential equations:

$$y' - y = v, \qquad \text{and} \qquad v' - 2v = 3x \tag{2.0.3}$$

because

$$y'' - y' = v' = 2v + 3x = 2y' - 2y + 3x, \tag{2.0.4}$$

which is the same as Equation 2.0.2. This suggests that Equation 2.0.2 can be solved by applying the techniques from the previous chapter. Proceeding along this line, we first find that

$$v(x) = C_1 e^{2x} - \tfrac{3}{2}x - \tfrac{3}{4}. \tag{2.0.5}$$

Therefore,

$$y' - y = C_1 e^{2x} - \tfrac{3}{2}x - \tfrac{3}{4}. \tag{2.0.6}$$

Again, applying the techniques from the previous chapter, we have that

$$y = C_1 e^{2x} + C_2 e^x + \tfrac{3}{2}x + \tfrac{9}{4}. \tag{2.0.7}$$

Note that the solution to this second-order ordinary differential equation contains *two* arbitrary constants.

[1] The proof of the existence and uniqueness of solutions to Equation 2.0.1 is beyond the scope of this book. See Ince, E. L., 1956: *Ordinary Differential Equations*. Dover Publications, Inc., Section 3.32.

• **Example 2.0.2**

In the case of linear, second-order ordinary differential equations, a similar technique, called *reduction in order*, provides a method for solving differential equations if we know one of its solutions.

Consider the second-order ordinary differential equation

$$x^2y'' - 5xy' + 9y = 0. \quad (\mathbf{2.0.8})$$

A quick check shows that $y_1(x) = x^3 \ln(x)$ is a solution of Equation 2.0.8. Let us now assume that the *general* solution can be written $y(x) = u(x)x^3 \ln(x)$. Then

$$y' = u'(x)x^3 \ln(x) + u(x) \left[3x^2 \ln(x) + x^2\right], \quad (\mathbf{2.0.9})$$

and

$$y'' = u''(x)x^3 \ln(x) + 2u'(x) \left[3x^2 \ln(x) + x^2\right] + u(x) \left[6x \ln(x) + 5x\right]. \quad (\mathbf{2.0.10})$$

Substitution of $y(x)$, $y'(x)$, and $y''(x)$ into (2.0.8) yields

$$x^5 \ln(x)u'' + \left[x^4 \ln(x) + 2x^4\right] u' = 0. \quad (\mathbf{2.0.11})$$

Setting $u' = w$, separation of variables leads to

$$\frac{w'}{w} = -\frac{1}{x} - \frac{2}{x \ln(x)}. \quad (\mathbf{2.0.12})$$

Note how our replacement of $u'(x)$ with $w(x)$ has reduced the second-order ordinary differential equation to a first-order one. Solving Equation 2.0.12, we find that

$$w(x) = u'(x) = -\frac{C_1}{x \ln^2(x)}, \quad (\mathbf{2.0.13})$$

and

$$u(x) = \frac{C_1}{\ln(x)} + C_2. \quad (\mathbf{2.0.14})$$

Because $y(x) = u(x)x^3 \ln(x)$, the complete solution is

$$y(x) = C_1x^3 + C_2x^3 \ln(x). \quad (\mathbf{2.0.15})$$

Substitution of Equation 2.0.15 into Equation 2.0.8 confirms that we have the correct solution.

We can verify our answer by using the symbolic toolbox in MATLAB. Typing the command:

```
dsolve('x*x*D2y-5*x*Dy+9*y=0','x')
```

yields

```
ans =
C1*x^3+C2*x^3*log(x)
```

□

In summary, we can reduce (in principle) any higher-order, linear ordinary differential equations into a system of first-order ordinary differential equations. This system of differential equations can then be solved using techniques from the previous chapter. In Chapter 3 we will pursue this idea further. Right now, however, we will introduce methods that allow us to find the solution in a more direct manner.

• Example 2.0.3

An *autonomous differential equation* is one where the independent variable does not appear explicitly. In certain cases we can reduce the order of the differential equation and then solve it.

Consider the autonomous ordinary differential equation

$$y'' = 2y^3. \tag{2.0.16}$$

The trick here is to note that

$$y'' = \frac{dv}{dx} = v\frac{dv}{dy} = 2y^3, \tag{2.0.17}$$

where $v = dy/dx$. Integrating both sides of Equation 2.0.17, we find that

$$v^2 = y^4 + C_1. \tag{2.0.18}$$

Solving for v,

$$\frac{dy}{dx} = v = \sqrt{C_1 + y^4}. \tag{2.0.19}$$

Integrating once more, we have the final result that

$$x + C_2 = \int \frac{dy}{\sqrt{C_1 + y^4}}. \tag{2.0.20}$$

Problems

For the following differential equations, use reduction of order to find a second solution. Can you obtain the general solution using `dsolve` in MATLAB?

1. $xy'' + 2y' = 0, \quad y_1(x) = 1$
2. $y'' + y' - 2y = 0, \quad y_1(x) = e^x$
3. $x^2y'' + 4xy' - 4y = 0, \quad y_1(x) = x$
4. $xy'' - (x+1)y' + y = 0, \quad y_1(x) = e^x$
5. $(2x - x^2)y'' + 2(x-1)y' - 2y = 0, \quad y_1(x) = x - 1$
6. $y'' + \tan(x)y' - 6\cot^2(x)y = 0, \quad y_1(x) = \sin^3(x)$
7. $4x^2y'' + 4xy' + (4x^2 - 1)y = 0, \quad y_1(x) = \cos(x)/\sqrt{x}$
8. $y'' + ay' + b(1 + ax - bx^2)y = 0, \quad y_1(x) = e^{-bx^2/2}$

Solve the following autonomous ordinary differential equations:

9. $yy'' = y'^2$
10. $y'' = 2yy', \quad y(0) = y'(0) = 1$

11. $yy'' = y' + y'^2$

12. $2yy'' = 1 + y'^2$

13. $y'' = e^{2y}, \quad y(0) = 0, y'(0) = 1$

14. $y''' = 3yy', \quad y(0) = y'(0) = 1, y''(0) = \frac{3}{2}$

15. Solve the nonlinear second-order ordinary differential equation

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

by (1) reducing it to the Bernoulli equation

$$\frac{dv}{dx} - \frac{v}{x} - \frac{v^2}{2} = 0, \qquad v(x) = u'(x),$$

(2) solving for $v(x)$, and finally (3) integrating $u' = v$ to find $u(x)$.

16. Consider the differential equation

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = 0, \qquad a_2(x) \neq 0.$$

Show that this ordinary differential equation can be rewritten

$$u'' + f(x)u = 0, \qquad f(x) = \frac{a_0(x)}{a_2(x)} - \frac{1}{4}\left[\frac{a_1(x)}{a_2(x)}\right]^2 - \frac{1}{2}\frac{d}{dx}\left[\frac{a_1(x)}{a_2(x)}\right],$$

using the substitution

$$y(x) = u(x)\exp\left[-\frac{1}{2}\int^x \frac{a_1(\xi)}{a_2(\xi)}\,d\xi\right].$$

2.1 HOMOGENEOUS LINEAR EQUATIONS WITH CONSTANT COEFFICIENTS

In our drive for more efficient methods to solve higher-order, linear, ordinary differential equations, let us examine the simplest possible case of a homogeneous differential equation with constant coefficients:

$$a_n\frac{d^ny}{dx^n} + a_{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_2y'' + a_1\frac{dy}{dx} + a_0y = 0. \tag{2.1.1}$$

Although we could explore Equation 2.1.1 in its most general form, we will begin by studying the second-order version, namely

$$ay'' + by' + cy = 0, \tag{2.1.2}$$

since it is the next step up the ladder in complexity from first-order ordinary differential equations.

Motivated by the fact that the solution to the first-order ordinary differential equation $y' + ay = 0$ is $y(x) = C_1e^{-ax}$, we make the educated guess that the solution to Equation 2.1.2 is $y(x) = Ae^{mx}$. Direct substitution into Equation 2.1.2 yields

$$\left(am^2 + bm + c\right)Ae^{mx} = 0. \tag{2.1.3}$$

The constant A cannot equal 0 because that would give $y(x) = 0$ and we would have a trivial solution. Furthermore, since $e^{mx} \neq 0$ for arbitrary x, Equation 2.1.3 simplifies to

$$am^2 + bm + c = 0. \tag{2.1.4}$$

Equation 2.1.4 is called the *auxiliary* or *characteristic equation*. At this point we must consider three separate cases.

• *Distinct real roots*

In this case the roots to Equation 2.1.4 are real and unequal. Let us denote these roots by $m = m_1$, and $m = m_2$. Thus, we have the two solutions:

$$y_1(x) = C_1 e^{m_1 x}, \qquad \text{and} \qquad y_2(x) = C_2 e^{m_2 x}. \tag{2.1.5}$$

We will now show that the most general solution to Equation 2.1.2 is

$$y(x) = C_1 e^{m_1 x} + C_2 e^{m_2 x}. \tag{2.1.6}$$

This result follows from the *principle of (linear) superposition*.

Theorem: *Let $y_1, y_2, \ldots, y_k$ be solutions of the homogeneous equation, Equation 2.1.1, on an interval I. Then the linear combination*

$$y(x) = C_1 y_1(x) + C_2 y_2(x) + \cdots + C_k y_k(x), \tag{2.1.7}$$

where C_i, $i = 1, 2, \ldots, k$, are arbitrary constants, is also a solution on the interval I.

Proof: We will prove this theorem for second-order ordinary differential equations; it is easily extended to higher orders. By the superposition principle, $y(x) = C_1 y_1(x) + C_2 y_2(x)$. Upon substitution into Equation 2.1.2, we have that

$$a\left(C_1 y_1'' + C_2 y_2''\right) + b\left(C_1 y_1' + C_2 y_2'\right) + c\left(C_1 y_1 + C_2 y_2\right) = 0. \tag{2.1.8}$$

Recombining the terms, we obtain

$$C_1\left(a y_1'' + b y_1' + c y_1\right) + C_2\left(a y_2'' + b y_2' + c y_2\right) = 0, \tag{2.1.9}$$

or

$$0C_1 + 0C_2 = 0. \tag{2.1.10}$$

□

• **Example 2.1.1**

A quick check shows that $y_1(x) = e^x$ and $y_2(x) = e^{-x}$ are two solutions of $y'' - y = 0$. Our theorem tells us that *any* linear combination of these solutions, such as $y(x) = 5e^x - 3e^{-x}$, is also a solution.

How about the converse? Is *every* solution to $y'' - y = 0$ a linear combination of $y_1(x)$ and $y_2(x)$? We will address this question shortly. □

• **Example 2.1.2**

Let us find the general solution to

$$y'' + 2y' - 15y = 0. \tag{2.1.11}$$

Assuming a solution of the form $y(x) = Ae^{mx}$, we have that

$$(m^2 + 2m - 15)Ae^{mx} = 0. \tag{2.1.12}$$

Because $A \neq 0$ and e^{mx} generally do not equal zero, we obtain the auxiliary or characteristic equation

$$m^2 + 2m - 15 = (m + 5)(m - 3) = 0. \tag{2.1.13}$$

Therefore, the general solution is

$$y(x) = C_1e^{3x} + C_2e^{-5x}. \tag{2.1.14}$$

□

• *Repeated real roots*

When $m = m_1 = m_2$, we have only the single exponential solution $y_1(x) = C_1e^{m_1x}$. To find the second solution we apply the reduction of order technique shown in Example 2.0.2. Performing the calculation, we find

$$y_2(x) = C_2e^{m_1x} \int \frac{e^{-bx/a}}{e^{2m_1x}}\, dx. \tag{2.1.15}$$

Since $m_1 = -b/(2a)$, the integral simplifies to $\int dx$ and

$$y(x) = C_1e^{m_1x} + C_2xe^{m_1x}. \tag{2.1.16}$$

• **Example 2.1.3**

Let us find the general solution to

$$y'' + 4y' + 4y = 0. \tag{2.1.17}$$

Here the auxiliary or characteristic equation is

$$m^2 + 4m + 4 = (m + 2)^2 = 0. \tag{2.1.18}$$

Therefore, the general solution is

$$y(x) = (C_1 + C_2x)e^{-2x}. \tag{2.1.19}$$

□

• *Complex conjugate roots*

When $b^2 - 4ac < 0$, the roots become the complex pair $m_1 = \alpha + i\beta$ and $m_2 = \alpha - \beta i$, where α and β are real and $i^2 = -1$. Therefore, the general solution is

$$y(x) = C_1 e^{(\alpha+i\beta)x} + C_2 e^{(\alpha-\beta i)x}. \tag{2.1.20}$$

Although Equation 2.1.20 is quite correct, most engineers prefer to work with real functions rather than complex exponentials. To this end, we apply Euler's formula[2] to eliminate $e^{i\beta x}$ and $e^{-i\beta x}$ since

$$e^{i\beta x} = \cos(\beta x) + i\sin(\beta x), \tag{2.1.21}$$

and

$$e^{-i\beta x} = \cos(\beta x) - i\sin(\beta x). \tag{2.1.22}$$

Therefore,

$$y(x) = C_1 e^{\alpha x}\left[\cos(\beta x) + i\sin(\beta x)\right] + C_2 e^{\alpha x}\left[\cos(\beta x) - i\sin(\beta x)\right] \tag{2.1.23}$$

$$= C_3 e^{\alpha x}\cos(\beta x) + C_4 e^{\alpha x}\sin(\beta x), \tag{2.1.24}$$

where $C_3 = C_1 + C_2$, and $C_4 = iC_1 - iC_2$.

• **Example 2.1.4**

Let us find the general solution to

$$y'' + 4y' + 5y = 0. \tag{2.1.25}$$

Here the auxiliary or characteristic equation is

$$m^2 + 4m + 5 = (m+2)^2 + 1 = 0, \tag{2.1.26}$$

or $m = -2 \pm i$. Therefore, the general solution is

$$y(x) = e^{-2x}[C_1\cos(x) + C_2\sin(x)]. \tag{2.1.27}$$

□

So far we have only dealt with second-order differential equations. When we turn to higher-order ordinary differential equations, similar considerations hold. In place of Equation 2.1.4, we now have the nth-degree polynomial equation

$$a_n m^n + a_{n-1}m^{n-1} + \cdots + a_2 m^2 + a_1 m + a_0 = 0 \tag{2.1.28}$$

for its auxiliary equation.

When we treated second-order ordinary differential equations, we were able to classify the roots to the auxiliary equation as distinct real roots, repeated roots, and complex roots. In the case of higher-order differential equations, such classifications are again useful

[2] If you are unfamiliar with Euler's formula, see Section 10.1.

although all three types may occur with the same equation. For example, the auxiliary equation

$$m^6 - m^5 + 2m^4 - 2m^3 + m^2 - m = 0 \qquad \textbf{(2.1.29)}$$

has the distinct roots $m = 0$ and $m = 1$ with the twice repeated, complex roots $m = \pm i$.

Although the possible combinations increase with higher-order differential equations, the solution technique remains the same. For each distinct real root $m = m_1$, we have a corresponding homogeneous solution $e^{m_1 x}$. For each complex pair $m = \alpha \pm \beta i$, we have the corresponding pair of homogeneous solutions $e^{\alpha x}\cos(\beta x)$ and $e^{\alpha x}\sin(\beta x)$. For a repeated root $m = m_1$ of multiplicity k, regardless of whether it is real or complex, we have either $e^{m_1 x}, xe^{m_1 x}, x^2 e^{m_1 x}, \ldots, x^k e^{m_1 x}$ in the case of real m_1 or

$$e^{\alpha x}\cos(\beta x), e^{\alpha x}\sin(\beta x), xe^{\alpha x}\cos(\beta x), xe^{\alpha x}\sin(\beta x),$$

$$x^2 e^{\alpha x}\cos(\beta x), x^2 e^{\alpha x}\sin(\beta x), \ldots, x^k e^{\alpha x}\cos(\beta x), x^k e^{\alpha x}\sin(\beta x)$$

in the case of complex roots $\alpha \pm \beta i$. For example, the general solution for the roots to Equation 2.1.29 is

$$y(x) = C_1 + C_2 e^x + C_3\cos(x) + C_4\sin(x) + C_5 x\cos(x) + C_6 x\sin(x). \qquad \textbf{(2.1.30)}$$

• Example 2.1.5

Let us find the general solution to

$$y''' + y' - 10y = 0. \qquad \textbf{(2.1.31)}$$

Here the auxiliary or characteristic equation is

$$m^3 + m - 10 = (m-2)(m^2 + 2m + 5) = (m-2)[(m+1)^2 + 4] = 0, \qquad \textbf{(2.1.32)}$$

or $m = -2$ and $m = -1 \pm 2i$. Therefore, the general solution is

$$y(x) = C_1 e^{-2x} + e^{-x}[C_2\cos(2x) + C_3\sin(2x)]. \qquad \textbf{(2.1.33)}$$

□

Having presented the technique for solving constant coefficient, linear, ordinary differential equations, an obvious question is: How do we know that we have captured all of the solutions? Before we can answer this question, we must introduce the concept of linear dependence.

A set of functions $f_1(x), f_2(x), \ldots, f_n(x)$ is said to be *linearly dependent* on an interval I if there exist constants $C_1, C_2, \ldots, C_n$, not all zero, such that

$$C_1 f_1(x) + C_2 f_2(x) + C_3 f_3(x) + \cdots + C_n f_n(x) = 0 \qquad \textbf{(2.1.34)}$$

for each x in the interval; otherwise, the set of functions is said to be *linearly independent*. This concept is easily understood when we have only two functions $f_1(x)$ and $f_2(x)$. If the functions are linearly dependent on an interval, then there exist constants C_1 and C_2 that are not both zero, where

$$C_1 f_1(x) + C_2 f_2(x) = 0 \qquad \textbf{(2.1.35)}$$

for every x in the interval. If $C_1 \neq 0$, then

$$f_1(x) = -\frac{C_2}{C_1} f_2(x). \tag{2.1.36}$$

In other words, if two functions are linearly dependent, then one is a constant multiple of the other. Conversely, two functions are linearly independent when neither is a constant multiple of the other on an interval.

• Example 2.1.6

Let us show that $f(x) = 2x$, $g(x) = 3x^2$, and $h(x) = 5x - 8x^2$ are linearly dependent on the real line.

To show this, we must choose three constants, C_1, C_2, and C_3, such that

$$C_1 f(x) + C_2 g(x) + C_3 h(x) = 0, \tag{2.1.37}$$

where not all of these constants are nonzero. A quick check shows that

$$15f(x) - 16g(x) - 6h(x) = 0. \tag{2.1.38}$$

Clearly, $f(x)$, $g(x)$, and $h(x)$ are linearly dependent. □

• Example 2.1.7

This example shows the importance of defining the interval on which a function is linearly dependent or independent. Consider the two functions $f(x) = x$ and $g(x) = |x|$. They are linearly dependent on the interval $(0, \infty)$ since $C_1 x + C_2|x| = C_1 x + C_2 x = 0$ is satisfied for any nonzero choice of C_1 and C_2 where $C_1 = -C_2$. What happens on the interval $(-\infty, 0)$? They are still linearly dependent but now $C_1 = C_2$. □

Although we could use the fundamental concept of linear independence to check and see whether a set of functions is linearly independent or not, the following theorem introduces a procedure that is very straightforward.

Theorem: Wronskian Test of Linear Independence

Suppose $f_1(x), f_2(x), \ldots, f_n(x)$ possess at least $n-1$ derivatives. If the determinant[3]

$$\begin{vmatrix} f_1 & f_2 & \cdots & f_n \\ f_1' & f_2' & \cdots & f_n' \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)} & f_2^{(n-1)} & \cdots & f_n^{(n-1)} \end{vmatrix}$$

is not zero for at least one point in the interval I, then the functions $f_1(x)$, $f_2(x)$, $\ldots$, $f_n(x)$ are linearly independent on the interval. The determinant in this theorem is denoted by $W[f_1(x), f_2(x), \ldots, f_n(x)]$ and is called the *Wronskian* of the functions.

[3] If you are unfamiliar with determinants, see Section 3.2.

Proof: We prove this theorem by contradiction when $n = 2$. Let us assume that $W[f_1(x_0), f_2(x_0)] \neq 0$ for some fixed x_0 in the interval I and that $f_1(x)$ and $f_2(x)$ are linearly dependent on the interval. Since the functions are linearly dependent, there exists C_1 and C_2, both not zero, for which

$$C_1 f_1(x) + C_2 f_2(x) = 0 \tag{2.1.39}$$

for every x in I. Differentiating Equation 2.1.39 gives

$$C_1 f_1'(x) + C_2 f_2'(x) = 0. \tag{2.1.40}$$

We may view Equation 2.1.39 and Equation 2.1.40 as a system of equations with C_1 and C_2 as the unknowns. Because the linear dependence of f_1 and f_2 implies that $C_1 \neq 0$ and/or $C_2 \neq 0$ for each x in the interval,

$$W[f_1(x), f_2(x)] = \begin{vmatrix} f_1 & f_2 \\ f_1' & f_2' \end{vmatrix} = 0 \tag{2.1.41}$$

for every x in I. This contradicts the assumption that $W[f_1(x_0), f_2(x_0)] \neq 0$ and f_1 and f_2 are linearly independent. □

• Example 2.1.8

Are the functions $f(x) = x$, $g(x) = xe^x$, and $h(x) = x^2e^x$ linearly dependent on the real line? To find out, we compute the Wronskian or

$$W[f(x), g(x), h(x)] = \begin{vmatrix} e^x & xe^x & x^2e^x \\ e^x & (x+1)e^x & (x^2+2x)e^x \\ e^x & (x+2)e^x & (x^2+4x+2)e^x \end{vmatrix} = e^{3x} \begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & 2x \\ 0 & 0 & 2 \end{vmatrix} = 2e^{3x} \neq 0. \tag{2.1.42}$$

Therefore, x, xe^x, and x^2e^x are linearly *independent*. □

Having introduced this concept of linear independence, we are now ready to address the question of how many linearly independent solutions a homogeneous linear equation has.

Theorem:

On any interval I over which an n-th order homogeneous linear differential equation is normal, the equation has n linearly independent solutions $y_1(x), y_2(x), \ldots, y_n(x)$ and any particular solution of the equation on I can be expressed as a linear combination of these linearly independent solutions.

Proof: Again for convenience and clarity we prove this theorem for the special case of $n = 2$. Let $y_1(x)$ and $y_2(x)$ denote solutions on I of Equation 2.1.2. We know that these solutions exist by the existence theorem and have the following values:

$$y_1(a) = 1, \quad y_2(a) = 0, \quad y_1'(a) = 0, \quad y_2'(a) = 1 \tag{2.1.43}$$

at some point a on I. To establish the linear independence of y_1 and y_2 we note that, if $C_1y_1(x) + C_2y_2(x) = 0$ holds identically on I, then $C_1y_1'(x) + C_2y_2'(x) = 0$ there too. Because $x = a$ lies in I, we have that

$$C_1 y_1(a) + C_2 y_2(a) = 0, \tag{2.1.44}$$

and

$$C_1 y_1'(a) + C_2 y_2'(a) = 0, \tag{2.1.45}$$

which yields $C_1 = C_2 = 0$ after substituting Equation 2.1.43. Hence, the solutions y_1 and y_2 are linearly independent.

To complete the proof we must now show that any particular solution of Equation 2.1.2 can be expressed as a linear combination of y_1 and y_2. Because y, y_1, and y_2 are all solutions of Equation 2.1.2 on I, so is the function

$$Y(x) = y(x) - y(a)y_1(x) - y'(a)y_2(x), \tag{2.1.46}$$

where $y(a)$ and $y'(a)$ are the values of the solution y and its derivative at $x = a$. Evaluating Y and Y' at $x = a$, we have that

$$Y(a) = y(a) - y(a)y_1(a) - y'(a)y_2(a) = y(a) - y(a) = 0, \tag{2.1.47}$$

and

$$Y'(a) = y'(a) - y(a)y_1'(a) - y'(a)y_2'(a) = y'(a) - y'(a) = 0. \tag{2.1.48}$$

Thus, Y is the trivial solution to Equation 2.1.2. Hence, for every x in I,

$$y(x) - y(a)y_1(x) - y'(a)y_2(x) = 0. \tag{2.1.49}$$

Solving Equation 2.1.49 for $y(x)$, we see that y is expressible as the linear combination

$$y(x) = y(a)y_1(x) + y'(a)y_2(x) \tag{2.1.50}$$

of y_1 and y_2, and the proof is complete for $n = 2$.

Problems

Find the general solution to the following differential equations. Check your general solution by using `dsolve` in MATLAB.

1. $y'' + 6y' + 5y = 0$
2. $y'' - 6y' + 10y = 0$
3. $y'' - 2y' + y = 0$
4. $y'' - 3y' + 2y = 0$
5. $y'' - 4y' + 8y = 0$
6. $y'' + 6y' + 9y = 0$
7. $y'' + 6y' - 40y = 0$
8. $y'' + 4y' + 5y = 0$
9. $y'' + 8y' + 25y = 0$
10. $4y'' - 12y' + 9y = 0$
11. $y'' + 8y' + 16y = 0$
12. $y''' + 4y'' = 0$
13. $y'''' + 4y'' = 0$
14. $y'''' + 2y''' + y'' = 0$
15. $y''' - 8y = 0$
16. $y'''' - 3y''' + 3y'' - y' = 0$

17. The simplest differential equation with "memory" — its past behavior affects the present — is

$$y' = -\frac{A}{2\tau}\int_{-\infty}^{t} e^{-(t-x)/\tau} y(x)\, dx.$$

Solve this integro-differential equation by differentiating it with respect to t to eliminate the integral.

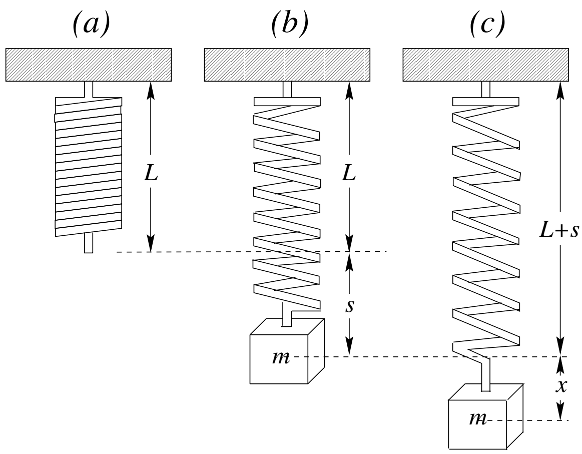


Figure 2.2.1: Various configurations of a mass/spring system. The spring alone has a length L, which increases to $L+s$ when the mass is attached. During simple harmonic motion, the length of the mass/spring system varies as $L+s+x$.

2.2 SIMPLE HARMONIC MOTION

Second-order, linear, ordinary differential equations often arise in mechanical or electrical problems. The purpose of this section is to illustrate how the techniques that we just derived may be applied to these problems.

We begin by considering the mass-spring system illustrated in Figure 2.2.1 where a mass m is attached to a flexible spring suspended from a rigid support. If there were no spring, then the mass would simply fall downward due to the gravitational force mg. Because there is no motion, the gravitational force must be balanced by an upward force due to the presence of the spring. This upward force is usually assumed to obey Hooke's law, which states that the restoring force is opposite to the direction of elongation and proportional to the amount of elongation. Mathematically the equilibrium condition can be expressed $mg = ks$.

Consider now what happens when we disturb this equilibrium. This may occur in one of two ways: We could move the mass either upward or downward and then release it. Another method would be to impart an initial velocity to the mass. In either case, the motion of the mass/spring system would be governed by Newton's second law, which states that the acceleration of the mass equals the imbalance of the forces. If we denote the downward displacement of the mass from its equilibrium position by positive x, then

$$m\frac{d^2x}{dt^2} = -k(s+x) + mg = -kx, \qquad \textbf{(2.2.1)}$$

since $ks = mg$. After dividing Equation 2.2.1 by the mass, we obtain the second-order differential equation

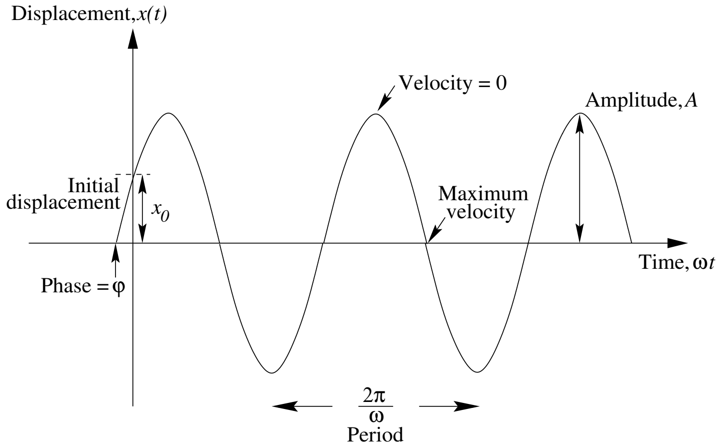
$$\frac{d^2x}{dt^2} + \frac{k}{m}x = 0, \qquad \textbf{(2.2.2)}$$

or

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \qquad \textbf{(2.2.3)}$$

where $\omega^2 = k/m$ and ω is the *circular frequency*. Equation 2.2.3 describes *simple harmonic motion* or *free undamped motion*. The two initial conditions associated with this differential

Summary of Simple Harmonic Motion



$$x(t) = \left(x_0^2 + \frac{v_0^2}{\omega^2}\right)^{1/2} \sin(\omega t + \varphi)$$

$$\varphi = \tan^{-1}\left(\frac{\omega x_0}{v_0}\right), \qquad v_0 = \text{initial velocity}$$

equation are

$$x(0) = \alpha, \qquad x'(0) = \beta. \tag{2.2.4}$$

The first condition gives the initial amount of displacement while the second condition specifies the initial velocity. If $\alpha > 0$ while $\beta < 0$, then the mass starts from a point below the equilibrium position with an initial upward velocity. On the other hand, if $\alpha < 0$ with $\beta = 0$ the mass is at rest when it is released $|\alpha|$ units above the equilibrium position. Similar considerations hold for other values of α and β.

To solve Equation 2.2.3, we note that the solutions of the auxiliary equation $m^2+\omega^2 = 0$ are the complex numbers $m_1 = \omega i$, and $m_2 = -\omega i$. Therefore, the general solution is

$$x(t) = A\cos(\omega t) + B\sin(\omega t). \tag{2.2.5}$$

The (natural) *period* of free vibrations is $T = 2\pi/\omega$ while the (natural) frequency is $f = 1/T = \omega/(2\pi)$.

• Example 2.2.1

Let us solve the initial-value problem

$$\frac{d^2x}{dt^2} + 4x = 0, \qquad x(0) = 10, \qquad x'(0) = 0. \tag{2.2.6}$$

The physical interpretation is that we have pulled the mass on a spring down 10 units *below* the equilibrium position and then release it from rest at $t = 0$. Here, $\omega = 2$ so that

$$x(t) = A\cos(2t) + B\sin(2t) \tag{2.2.7}$$

from Equation 2.2.5.

Because $x(0) = 10$, we find that

$$x(0) = 10 = A \cdot 1 + B \cdot 0 \tag{2.2.8}$$

so that $A = 10$. Next, we note that

$$\frac{dx}{dt} = -20\sin(2t) + 2B\cos(2t). \tag{2.2.9}$$

Therefore, at $t = 0$,

$$x'(0) = 0 = -20 \cdot 0 + 2B \cdot 1 \tag{2.2.10}$$

and $B = 0$. Thus, the equation of motion is $x(t) = 10\cos(2t)$.

What is the physical interpretation of our equation of motion? Once the system is set into motion, it stays in motion with the mass oscillating back and forth 10 units above and below the equilibrium position $x = 0$. The period of oscillation is $2\pi/2 = \pi$ units of time.□

• Example 2.2.2

A weight of 45 N stretches a spring 5 cm. At time $t = 0$, the weight is released from its equilibrium position with an upward velocity of 28 cm s^{-1}. Determine the displacement $x(t)$ that describes the subsequent free motion.

From Hooke's law,

$$F = mg = 45\,\text{N} = k \times 5\,\text{cm} \tag{2.2.11}$$

so that $k = 9$ N cm^{-1}. Therefore, the differential equation is

$$\frac{d^2x}{dt^2} + 196\,\text{s}^{-2}x = 0. \tag{2.2.12}$$

The initial displacement and initial velocity are $x(0) = 0$ cm and $x'(0) = -28$ cm s^{-1}. The negative sign in the initial velocity reflects the fact that the weight has an initial velocity in the negative or upward direction.

Because $\omega^2 = 196$ s^{-2} or $\omega = 14$ s^{-1}, the general solution to the differential equation is

$$x(t) = A\cos(14\,\text{s}^{-1}t) + B\sin(14\,\text{s}^{-1}t). \tag{2.2.13}$$

Substituting for the initial displacement $x(0)$ in Equation 2.2.13, we find that

$$x(0) = 0\,\text{cm} = A \cdot 1 + B \cdot 0, \tag{2.2.14}$$

and $A = 0$ cm. Therefore,

$$x(t) = B\sin(14\,\text{s}^{-1}t) \tag{2.2.15}$$

and

$$x'(t) = 14\,\text{s}^{-1}B\cos(14\,\text{s}^{-1}t). \tag{2.2.16}$$

Substituting for the initial velocity,

$$x'(0) = -28\,\text{cm}\,\text{s}^{-1} = 14\,\text{s}^{-1}B, \tag{2.2.17}$$

and $B = -2$ cm. Thus the equation of motion is

$$x(t) = -2\,\text{cm}\,\sin(14\,\text{s}^{-1}t). \tag{2.2.18}$$

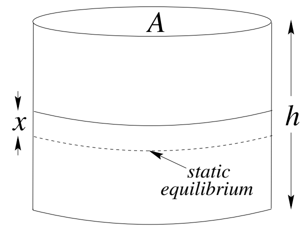


Figure 2.2.2: Schematic of a floating body partially submerged in pure water.

□

• **Example 2.2.3: Vibration of floating bodies**

Consider a solid cylinder of radius a that is partially submerged in a bath of pure water as shown in Figure 2.2.2. Let us find the motion of this cylinder in the vertical direction assuming that it remains in an upright position.

If the displacement of the cylinder from its static equilibrium position is x, the weight of water displaced equals $Ag\rho_w x$, where ρ_w is the density of the water and g is the gravitational acceleration. This is the restoring force according to the Archimedes principle. The mass of the cylinder is $Ah\rho$, where ρ is the density of cylinder. From second Newton's law, the equation of motion is

$$\rho A h x'' + A g \rho_w x = 0, \tag{2.2.19}$$

or

$$x'' + \frac{\rho_w g}{\rho h} x = 0. \tag{2.2.20}$$

From Equation 2.2.20 we see that the cylinder will oscillate about its static equilibrium position $x = 0$ with a frequency of

$$\omega = \left(\frac{\rho_w g}{\rho h} \right)^{1/2}. \tag{2.2.21}$$

□

When both A and B are both nonzero, it is often useful to rewrite the homogeneous solution, Equation 2.2.5, as

$$x(t) = C \sin(\omega t + \varphi) \tag{2.2.22}$$

to highlight the amplitude and phase of the oscillation. Upon employing the trigonometric angle-sum formula, Equation 2.2.22 can be rewritten

$$x(t) = C \sin(\omega t) \cos(\varphi) + C \cos(\omega t) \sin(\varphi) = A \cos(\omega t) + B \sin(\omega t). \tag{2.2.23}$$

From Equation 2.2.23, we see that $A = C \sin(\varphi)$ and $B = C \cos(\varphi)$. Therefore,

$$A^2 + B^2 = C^2 \sin^2(\varphi) + C^2 \cos^2(\varphi) = C^2, \tag{2.2.24}$$

and $C = \sqrt{A^2 + B^2}$. Similarly, $\tan(\varphi) = A/B$. Because the tangent is positive in both the first and third quadrants and negative in both the second and fourth quadrants, there are two possible choices for φ. The proper value of φ satisfies the equations $A = C\sin(\varphi)$ and $B = C\cos(\varphi)$.

If we prefer the amplitude/phase solution

$$x(t) = C\cos(\omega t - \varphi), \tag{2.2.25}$$

we now have

$$x(t) = C\cos(\omega t)\cos(\varphi) + C\sin(\omega t)\sin(\varphi) = A\cos(\omega t) + B\sin(\omega t). \tag{2.2.26}$$

Consequently, $A = C\cos(\varphi)$ and $B = C\sin(\varphi)$. Once again, we obtain $C = \sqrt{A^2 + B^2}$. On the other hand, $\tan(\varphi) = B/A$.

Problems

Solve the following initial-value problems and write their solutions in terms of amplitude and phase:

1. $x'' + 25x = 0, \qquad x(0) = 10, \qquad x'(0) = -10$

2. $4x'' + 9x = 0, \qquad x(0) = 2\pi, \qquad x'(0) = 3\pi$

3. $x'' + \pi^2 x = 0, \qquad x(0) = 1, \qquad x'(0) = \pi\sqrt{3}$

4. A 4-kg mass is suspended from a 100 N/m spring. The mass is set in motion by giving it an initial downward velocity of 5 m/s from its equilibrium position. Find the displacement as a function of time.

5. A spring hangs vertically. A weight of mass M kg stretches it L m. This weight is removed. A body weighing m kg is then attached and allowed to come to rest. It is then pulled down s_0 m and released with a velocity v_0. Find the displacement of the body from its point of rest and its velocity at any time t.

6. A particle of mass m moving in a straight line is *repelled* from the origin by a force F. (a) If the force is proportional to the distance from the origin, find the position of the particle as a function of time. (b) If the initial velocity of the particle is $a\sqrt{k}$, where k is the proportionality constant and a is the distance from the origin, find the position of the particle as a function of time. What happens if $m < 1$ and $m = 1$?

2.3 DAMPED HARMONIC MOTION

Free harmonic motion is unrealistic because there are always frictional forces that act to retard motion. In mechanics, the drag is often modeled as a resistance that is proportional to the instantaneous velocity. Adopting this resistance law, it follows from Newton's second law that the harmonic oscillator is governed by

$$m\frac{d^2x}{dt^2} = -kx - \beta\frac{dx}{dt}, \tag{2.3.1}$$

where β is a positive *damping constant*. The negative sign is necessary since this resistance acts in a direction opposite to the motion.

Dividing Equation 2.3.1 by the mass m, we obtain the differential equation of *free damped motion*,

$$\frac{d^2x}{dt^2} + \frac{\beta}{m}\frac{dx}{dt} + \frac{k}{m}x = 0, \tag{2.3.2}$$

or

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + \omega^2 x = 0. \tag{2.3.3}$$

We have written 2λ rather than just λ because it simplifies future computations. The auxiliary equation is $m^2 + 2\lambda m + \omega^2 = 0$, which has the roots

$$m_1 = -\lambda + \sqrt{\lambda^2 - \omega^2}, \qquad \text{and} \qquad m_2 = -\lambda - \sqrt{\lambda^2 - \omega^2}. \tag{2.3.4}$$

From Equation 2.3.4 we see that there are three possible cases which depend on the algebraic sign of $\lambda^2 - \omega^2$. Because all of the solutions contain the damping factor $e^{-\lambda t}$, $\lambda > 0$, $x(t)$ vanishes as $t \to \infty$.

- *Case I:* $\lambda > \omega$

Here the system is *overdamped* because the damping coefficient β is large compared to the spring constant k. The corresponding solution is

$$x(t) = Ae^{m_1 t} + Be^{m_2 t}, \tag{2.3.5}$$

or

$$x(t) = e^{-\lambda t}\left(Ae^{t\sqrt{\lambda^2-\omega^2}} + Be^{-t\sqrt{\lambda^2-\omega^2}}\right). \tag{2.3.6}$$

In this case the motion is smooth and nonoscillatory.

- *Case II:* $\lambda = \omega$

The system is *critically damped* because any slight decrease in the damping force would result in oscillatory motion. The general solution is

$$x(t) = Ae^{m_1 t} + Bte^{m_1 t}, \tag{2.3.7}$$

or

$$x(t) = e^{-\lambda t}(A + Bt). \tag{2.3.8}$$

The motion is quite similar to that of an overdamped system.

- *Case III:* $\lambda < \omega$

In this case the system is *underdamped* because the damping coefficient is small compared to the spring constant. The roots m_1 and m_2 are complex:

$$m_1 = -\lambda + i\sqrt{\omega^2 - \lambda^2}, \qquad \text{and} \qquad m_2 = -\lambda - i\sqrt{\omega^2 - \lambda^2}. \tag{2.3.9}$$

The general solution now becomes

$$x(t) = e^{-\lambda t}\left[A\cos\left(t\sqrt{\omega^2 - \lambda^2}\right) + B\sin\left(t\sqrt{\omega^2 - \lambda^2}\right)\right]. \tag{2.3.10}$$

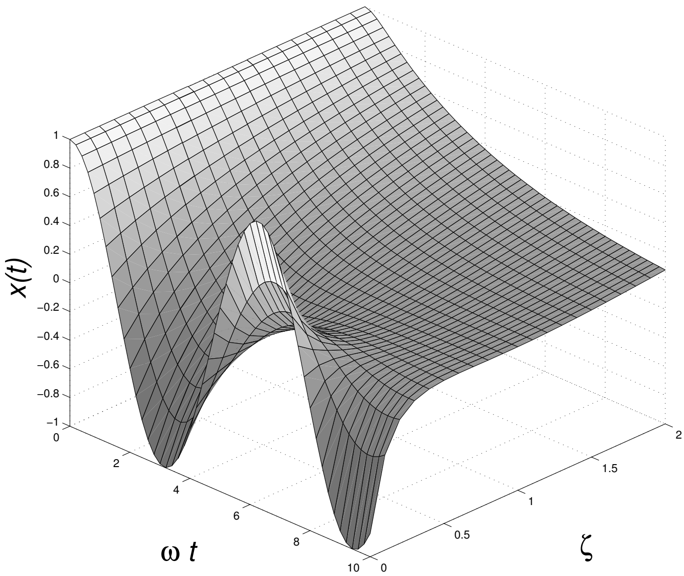


Figure 2.3.1: The displacement $x(t)$ of a damped harmonic oscillator as a function of time and $\zeta = \lambda/\omega$.

Equation 2.3.10 describes oscillatory motion that decays as $e^{-\lambda t}$. Equations 2.3.6, 2.3.8, and 2.3.10 are illustrated in Figure 2.3.1 when the initial conditions are $x(0) = 1$ and $x'(0) = 0$.

Just as we could write the solution for the simple harmonic motion in the amplitude/phase format, we can write any damped solution Equation 2.3.10 in the alternative form

$$x(t) = Ce^{-\lambda t} \sin\left(t\sqrt{\omega^2 - \lambda^2} + \varphi\right), \tag{2.3.11}$$

where $C = \sqrt{A^2 + B^2}$ and the phase angle φ is given by $\tan(\varphi) = A/B$ such that $A = C\sin(\varphi)$ and $B = C\cos(\varphi)$. The coefficient $Ce^{-\lambda t}$ is sometimes called the *damped coefficient* of vibrations. Because Equation 2.3.11 is *not* a periodic function, the quantity $2\pi/\sqrt{\omega^2 - \lambda^2}$ is called the *quasi period* and $\sqrt{\omega^2 - \lambda^2}$ is the *quasi frequency*. The quasi period is the time interval between two successive maxima of $x(t)$.

• Example 2.3.1

A body with mass $m = \frac{1}{2}$ kg is attached to the end of a spring that is stretched 2 m by a force of 100 N. Furthermore, there is also attached a dashpot[4] that provides 6 N of resistance for each m/s of velocity. If the mass is set in motion by further stretching the spring $\frac{1}{2}$ m and giving it an upward velocity of 10 m/s, let us find the subsequent motion.

We begin by first computing the constants. The spring constant is $k = (100\text{ N})/(2\text{ m}) = 50$ N/m. Therefore, the differential equation is

$$\tfrac{1}{2}x'' + 6x' + 50x = 0 \tag{2.3.12}$$

with $x(0) = \frac{1}{2}$ m and $x'(0) = -10$ m/s. Here the units of $x(t)$ are meters. The characteristic or auxiliary equation is

$$m^2 + 12m + 100 = (m + 6)^2 + 64 = 0, \tag{2.3.13}$$

[4] A mechanical device — usually a piston that slides within a liquid-filled cylinder — used to damp the vibration or control the motion of a mechanism to which is attached.

Review of the Solution of the Underdamped Homogeneous Oscillator Problem

$mx'' + \beta x' + kx = 0$ subject to $x(0) = x_0$, $x'(0) = v_0$ has the solution

$$x(t) = Ae^{-\lambda t}\sin(\omega_d t + \varphi),$$

where

$$\omega = \sqrt{k/m} \text{ is the undamped natural frequency,}$$
$$\lambda = \beta/(2m) \text{ is the damping factor,}$$
$$\omega_d = \sqrt{\omega^2 - \lambda^2} \text{ is the damped natural frequency,}$$

and the constants A and φ are determined by

$$A = \sqrt{x_0^2 + \left(\frac{v_0 + \lambda x_0}{\omega_d}\right)^2}$$

and

$$\varphi = \tan^{-1}\left(\frac{x_0\omega_d}{v_0 + \lambda x_0}\right).$$

or $m = -6 \pm 8i$. Therefore, we have an underdamped harmonic oscillator and the general solution is

$$x(t) = e^{-6t}\left[A\cos(8t) + B\sin(8t)\right]. \quad \textbf{(2.3.14)}$$

Consequently, each cycle takes $2\pi/8 = 0.79$ second. This is longer than the 0.63 second that would occur if the system were undamped.

From the initial conditions,

$$x(0) = A = \tfrac{1}{2}, \quad \text{and} \quad x'(0) = -10 = -6A + 8B. \quad \textbf{(2.3.15)}$$

Therefore, $A = \frac{1}{2}$ and $B = -\frac{7}{8}$. Consequently,

$$x(t) = e^{-6t}\left[\tfrac{1}{2}\cos(8t) - \tfrac{7}{8}\sin(8t)\right] = \frac{\sqrt{65}}{8}e^{-6t}\cos(8t + 2.62244). \quad \textbf{(2.3.16)}$$

□

• Example 2.3.2: Design of a wind vane

In its simplest form a wind vane is a flat plate or airfoil that can rotate about a vertical shaft. See Figure 2.3.2. In static equilibrium it points into the wind. There is usually a counterweight to balance the vane about the vertical shaft.

A vane uses a combination of the lift and drag forces on the vane to align itself with the wind. As the wind shifts direction from θ_0 to the new direction θ_i, the direction θ in which the vane currently points is governed by the equation of motion[5]

$$I\frac{d^2\theta}{dt^2} + \frac{NR}{V}\frac{d\theta}{dt} = N(\theta_i - \theta), \quad \textbf{(2.3.17)}$$

[5] For a derivation of Equation 2.3.12 and Equation 2.3.13, see subsection 2 of Section 3 in Barthelt, H. P., and G. H. Ruppersberg, 1957: Die mechanische Windfahne, eine theoretische und experimentelle Untersuchung. *Beitr. Phys. Atmos.*, **29**, 154–185.

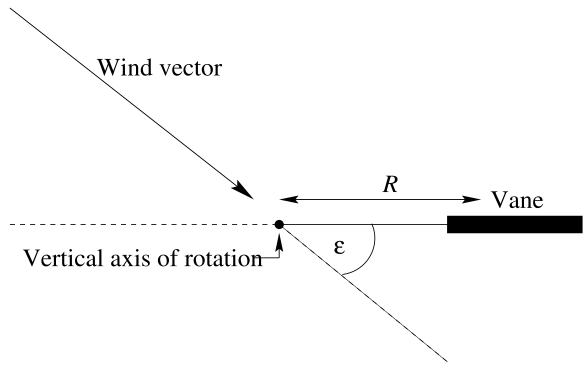


Figure 2.3.2: Schematic of a wind vane. The counterbalance is not shown.

where I is the vane's moment of inertia, N is the aerodynamic torque per unit angle, and R is the distance from the axis of rotation to the effective center of the aerodynamic force on the vane. The aerodynamic torque is given by

$$N = \tfrac{1}{2} C_L \rho A V^2 R, \tag{2.3.18}$$

where C_L is the lift coefficient, ρ is the air density, A is the vane area, and V is the wind speed.

Dividing Equation 2.3.17 by I, we obtain the second-order ordinary differential equation

$$\frac{d^2(\theta - \theta_i)}{dt^2} + \frac{NR}{IV}\frac{d(\theta - \theta_i)}{dt} + \frac{N}{I}(\theta - \theta_i) = 0. \tag{2.3.19}$$

The solution to Equation 2.3.19 is

$$\theta - \theta_i = A \exp\left(-\frac{NRt}{2IV}\right)\cos(\omega t + \varphi), \tag{2.3.20}$$

where

$$\omega^2 = \frac{N}{I} - \frac{N^2R^2}{4I^2V^2}, \tag{2.3.21}$$

and A and φ are the two arbitrary constants that would be determined by presently unspecified initial conditions. Consequently an ideal wind vane is a damped harmonic oscillator where the wind torque should be large and its moment of inertia should be small.

Problems

For the following values of m, β, and k, find the position $x(t)$ of a damped oscillator for the given initial conditions:

1. $m = \frac{1}{2}, \quad \beta = 3, \quad k = 4, \quad x(0) = 2, \quad x'(0) = 0$

2. $m = 1, \quad \beta = 10, \quad k = 125, \quad x(0) = 3, \quad x'(0) = 25$

3. $m = 4, \quad \beta = 20, \quad k = 169, \quad x(0) = 4, \quad x'(0) = 16$

4. For a fixed value of λ/ω, what is the minimum number of cycles required to produce a reduction of at least 50% in the maxima of an underdamped oscillator?

5. For what values of c does $x'' + cx' + 4x = 0$ have critically damped solutions?

6. For what values of c are the motions governed by $4x'' + cx' + 9x = 0$ (a) overdamped, (b) underdamped, and (c) critically damped?

7. For an overdamped mass-spring system, prove that the mass can pass through its equilibrium position $x = 0$ at most once.

2.4 METHOD OF UNDETERMINED COEFFICIENTS

Homogeneous ordinary differential equations become nonhomogeneous when the right side of Equation 2.0.1 is nonzero. How does this case differ from the homogeneous one that we have treated so far?

To answer this question, let us begin by introducing a function $y_p(x)$ — called a *particular solution* — whose only requirement is that it satisfies the differential equation

$$a_n(x)\frac{d^n y_p}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y_p}{dx^{n-1}} + \cdots + a_1(x)\frac{dy_p}{dx} + a_0(x)y_p = f(x). \qquad \textbf{(2.4.1)}$$

Then, by direct substitution, it can be seen that the general solution to any nonhomogeneous, linear, ordinary differential equation is

$$y(x) = y_H(x) + y_p(x), \qquad \textbf{(2.4.2)}$$

where $y_H(x)$ — the *homogeneous* or *complementary solution* — satisfies

$$a_n(x)\frac{d^n y_H}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y_H}{dx^{n-1}} + \cdots + a_1(x)\frac{dy_H}{dx} + a_0(x)y_H = 0. \qquad \textbf{(2.4.3)}$$

Why have we introduced this complementary solution - because the particular solution already satisfies the ordinary differential equation. The purpose of the complementary solution is to introduce the arbitrary constants that any general solution of an ordinary differential equation must have. Thus, because we already know how to find $y_H(x)$, we must only invent a method for finding the particular solution to have our general solution.

• Example 2.4.1

Let us illustrate this technique with the second-order, linear, nonhomogeneous ordinary differential equation

$$y'' - 4y' + 4y = 2e^{2x} + 4x - 12. \qquad \textbf{(2.4.4)}$$

Taking $y(x) = y_H(x) + y_p(x)$, direction substitution yields

$$y_H'' + y_p'' - 4(y_H' + y_p') + 4(y_H + y_p) = 2e^{2x} + 4x - 12. \qquad \textbf{(2.4.5)}$$

If we now require that the particular solution $y_p(x)$ satisfies the differential equation

$$y_p'' - 4y_p' + 4y_p = 2e^{2x} + 4x - 12, \qquad \textbf{(2.4.6)}$$

Equation 2.4.5 simplifies to the homogeneous ordinary differential equation

$$y_H'' - 4y_H' + 4y_H = 0. \qquad \textbf{(2.4.7)}$$

A quick check[6] shows that the particular solution to Equation 2.4.6 is $y_p(x) = x^2e^{2x} + x - 2$. Using techniques from the previous section, the complementary solution is $y_H(x) = C_1e^{2x} + C_2xe^{2x}$. □

In general, finding $y_p(x)$ is a formidable task. In the case of constant coefficients, several techniques have been developed. The most commonly employed technique is called the *method of undetermined coefficients*, which is used with linear, constant coefficient, ordinary differential equations when $f(x)$ is a constant, a polynomial, an exponential function $e^{\alpha x}$, $\sin(\beta x)$, $\cos(\beta x)$, or finite sum and products of these functions. Thus, this technique applies when the function $f(x)$ equals $e^x \sin(x) - (3x - 2)e^{-2x}$ but not when it equals $\ln(x)$.

Why does this technique work? The reason lies in the set of functions that we have allowed to be included in $f(x)$. They enjoy the remarkable property that derivatives of their sums and products yield sums and products that are also constants, polynomials, exponentials, sines, and cosines. Because a linear combination of derivatives such as $ay_p'' + by_p' + cy_p$ must equal $f(x)$, it seems reasonable to assume that $y_p(x)$ has the same form as $f(x)$. The following examples show that our conjecture is correct.

• Example 2.4.2

Let us illustrate the method of undetermined coefficients by finding the particular solution to

$$y'' - 2y' + y = x + \sin(x) \qquad (\mathbf{2.4.8})$$

by the method of undetermined coefficients.

From the form of the right side of Equation 2.4.8, we guess the particular solution

$$y_p(x) = Ax + B + C\sin(x) + D\cos(x). \qquad (\mathbf{2.4.9})$$

Therefore,

$$y_p'(x) = A + C\cos(x) - D\sin(x), \qquad (\mathbf{2.4.10})$$

and

$$y_p''(x) = -C\sin(x) - D\cos(x). \qquad (\mathbf{2.4.11})$$

Substituting into Equation 2.4.8, we find that

$$y_p'' - 2y_p' + y_p = Ax + B - 2A - 2C\cos(x) + 2D\sin(x) = x + \sin(x). \qquad (\mathbf{2.4.12})$$

Since Equation 2.4.12 must be true for all x, the constant terms must sum to zero or $B - 2A = 0$. Similarly, all of the terms involving the polynomial x must balance, yielding $A = 1$ and $B = 2A = 2$. Turning to the trigonometric terms, the coefficients of $\sin(x)$ and $\cos(x)$ give $2D = 1$ and $-2C = 0$, respectively. Therefore, the particular solution is

$$y_p(x) = x + 2 + \tfrac{1}{2}\cos(x), \qquad (\mathbf{2.4.13})$$

and the general solution is

$$y(x) = y_H(x) + y_p(x) = C_1e^x + C_2xe^x + x + 2 + \tfrac{1}{2}\cos(x). \qquad (\mathbf{2.4.14})$$

[6] We will show how $y_p(x)$ was obtained momentarily.

We can verify our result by using the symbolic toolbox in MATLAB. Typing the command:

```
dsolve('D2y-2*Dy+y=x+sin(x)','x')
```

yields

```
ans =
x+2+1/2*cos(x)+C1*exp(x)+C2*exp(x)*x
```

□

• Example 2.4.3

Let us find the particular solution to

$$y'' + y' - 2y = xe^x \qquad (2.4.15)$$

by the method of undetermined coefficients.

From the form of the right side of Equation 2.4.15, we guess the particular solution

$$y_p(x) = Axe^x + Be^x. \qquad (2.4.16)$$

Therefore,

$$y_p'(x) = Axe^x + Ae^x + Be^x, \qquad (2.4.17)$$

and

$$y_p''(x) = Axe^x + 2Ae^x + Be^x. \qquad (2.4.18)$$

Substituting into Equation 2.4.15, we find that

$$3Ae^x = xe^x. \qquad (2.4.19)$$

Clearly we cannot choose a constant A such that Equation 2.4.19 is satisfied. What went wrong?

To understand why, let us find the homogeneous or complementary solution to Equation 2.4.15; it is

$$y_H(x) = C_1 e^{-2x} + C_2 e^x. \qquad (2.4.20)$$

Therefore, one of the assumed particular solutions, Be^x, is also a homogeneous solution and cannot possibly give a nonzero left side when substituted into the differential equation. Consequently, it would appear that the method of undetermined coefficients does not work when one of the terms on the right side is also a homogeneous solution.

Before we give up, let us recall that we had a similar situation in the case of linear homogeneous second-order ordinary differential equations when the roots from the auxiliary equation were equal. There we found one of the homogeneous solutions was $e^{m_1 x}$. We eventually found that the second solution was $xe^{m_1 x}$. Could such a solution work here? Let us try.

We begin by modifying Equation 2.4.16 by multiplying it by x. Thus, our new guess for the particular solution reads

$$y_p(x) = Ax^2e^x + Bxe^x. \qquad (2.4.21)$$

Then,

$$y_p' = Ax^2e^x + 2Axe^x + Bxe^x + Be^x, \qquad (2.4.22)$$

and

$$y_p'' = Ax^2e^x + 4Axe^x + 2Ae^x + Bxe^x + 2Be^x. \qquad \textbf{(2.4.23)}$$

Substituting Equation 2.4.21 into Equation 2.4.15 gives

$$y_p'' + y_p' - 2y_p = 6Axe^x + 2Ae^x + 3Be^x = xe^x. \qquad \textbf{(2.4.24)}$$

Grouping together terms that vary as xe^x, we find that $6A = 1$. Similarly, terms that vary as e^x yield $2A + 3B = 0$. Therefore,

$$y_p(x) = \tfrac{1}{6}x^2e^x - \tfrac{1}{9}xe^x, \qquad \textbf{(2.4.25)}$$

so that the general solution is

$$y(x) = y_H(x) + y_p(x) = C_1e^{-2x} + C_2e^x + \tfrac{1}{6}x^2e^x - \tfrac{1}{9}xe^x. \qquad \textbf{(2.4.26)}$$

□

In summary, the method of finding particular solutions to higher-order ordinary differential equations by the method of undetermined coefficients is as follows:

- **Step 1:** Find the homogeneous solution to the differential equation.
- **Step 2:** Make an initial guess at the particular solution. The form of $y_p(x)$ is a linear combination of all linearly independent functions that are generated by repeated differentiations of $f(x)$.
- **Step 3:** If any of the terms in $y_p(x)$ given in Step 2 duplicate any of the homogeneous solutions, then that particular term in $y_p(x)$ must be multiplied by x^n, where n is the smallest positive integer that eliminates the duplication.

• **Example 2.4.4**

Let us apply the method of undetermined coefficients to solve

$$y'' + y = \sin(x) - e^{3x}\cos(5x). \qquad \textbf{(2.4.27)}$$

We begin by first finding the solution to the homogeneous version of Equation 2.4.27:

$$y_H'' + y_H = 0. \qquad \textbf{(2.4.28)}$$

Its solution is

$$y_H(x) = A\cos(x) + B\sin(x). \qquad \textbf{(2.4.29)}$$

To find the particular solution we examine the right side of Equation 2.4.27 or

$$f(x) = \sin(x) - e^{3x}\cos(5x). \qquad \textbf{(2.4.30)}$$

Taking a few derivatives of $f(x)$, we find that

$$f'(x) = \cos(x) - 3e^{3x}\cos(5x) + 5e^{3x}\sin(5x), \qquad \textbf{(2.4.31)}$$

$$f''(x) = -\sin(x) - 9e^{3x}\cos(5x) + 30e^{3x}\sin(5x) + 25e^{3x}\cos(5x), \qquad \textbf{(2.4.32)}$$

and so forth. Therefore, our guess at the particular solution is

$$y_p(x) = Cx\sin(x) + Dx\cos(x) + Ee^{3x}\cos(5x) + Fe^{3x}\sin(5x). \qquad \textbf{(2.4.33)}$$

Why have we chosen $x\sin(x)$ and $x\cos(x)$ rather than $\sin(x)$ and $\cos(x)$? Because $\sin(x)$ and $\cos(x)$ are homogeneous solutions to Equation 2.4.27, we must multiply them by a power of x.

Since

$$y_p''(x) = 2C\cos(x) - Cx\sin(x) - 2D\sin(x) - Dx\cos(x)$$
$$+ (30F - 16E)e^{3x}\cos(5x) - (30E + 16F)e^{3x}\sin(5x), \qquad \textbf{(2.4.34)}$$

$$y_p'' + y_p = 2C\cos(x) - 2D\sin(x)$$
$$+ (30F - 15E)e^{3x}\cos(5x) - (30E + 15F)e^{3x}\sin(5x) \qquad \textbf{(2.4.35)}$$
$$= \sin(x) - e^{3x}\cos(5x). \qquad \textbf{(2.4.36)}$$

Therefore, $2C = 0$, $-2D = 1$, $30F - 15E = -1$, and $30E + 15F = 0$. Solving this system of equations yields $C = 0$, $D = -\frac{1}{2}$, $E = \frac{1}{75}$, and $F = -\frac{2}{75}$. Thus, the general solution is

$$y(x) = A\cos(x) + B\sin(x) - \tfrac{1}{2}x\cos(x) + \tfrac{1}{75}e^{3x}[\cos(5x) - 2\sin(5x)]. \qquad \textbf{(2.4.37)}$$

Problems

Use the method of undetermined coefficients to find the general solution of the following differential equations. Verify your solution by using `dsolve` in MATLAB.

1. $y'' + 4y' + 3y = x + 1$
2. $y'' - y = e^x - 2e^{-2x}$
3. $y'' + 2y' + 2y = 2x^2 + 2x + 4$
4. $y'' + y' = x^2 + x$
5. $y'' + 2y' = 2x + 5 - e^{-2x}$
6. $y'' - 4y' + 4y = (x + 1)e^{2x}$
7. $y'' + 4y' + 4y = xe^x$
8. $y'' - 4y = 4\sinh(2x)$
9. $y'' + 9y = x\cos(3x)$
10. $y'' + y = \sin(x) + x\cos(x)$

11. Solve

$$y'' + 2ay' = \sin^2(\omega x), \qquad y(0) = y'(0) = 0,$$

by (a) the method of undetermined coefficients and (b) integrating the ordinary differential equation so that it reduces to

$$y' + 2ay = \frac{x}{2} - \frac{\sin(2ax)}{4a},$$

and then using the techniques from the previous chapter to solve this first-order ordinary differential equation.

2.5 FORCED HARMONIC MOTION

Let us now consider the situation when an external force $f(t)$ acts on a vibrating mass on a spring. For example, $f(t)$ could represent a driving force that periodically raises and lowers the support of the spring. The inclusion of $f(t)$ in the formulation of Newton's second law yields the differential equation

$$m\frac{d^2x}{dt^2} = -kx - \beta\frac{dx}{dt} + f(t), \tag{2.5.1}$$

$$\frac{d^2x}{dt^2} + \frac{\beta}{m}\frac{dx}{dt} + \frac{k}{m}x = \frac{f(t)}{m}, \tag{2.5.2}$$

or

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + \omega^2 x = F(t), \tag{2.5.3}$$

where $F(t) = f(t)/m$, $2\lambda = \beta/m$, and $\omega^2 = k/m$. To solve this nonhomogeneous equation we will use the method of undetermined coefficients.

• Example 2.5.1

Let us find the solution to the nonhomogeneous differential equation

$$y'' + 2y' + y = 2\sin(t), \tag{2.5.4}$$

subject to the initial conditions $y(0) = 2$ and $y'(0) = 1$.

The homogeneous solution is easily found and equals

$$y_H(t) = Ae^{-t} + Bte^{-t}. \tag{2.5.5}$$

From the method of undetermined coefficients, we guess that the particular solution is

$$y_p(t) = C\cos(t) + D\sin(t), \tag{2.5.6}$$

so that

$$y_p'(t) = -C\sin(t) + D\cos(t), \tag{2.5.7}$$

and

$$y_p''(t) = -C\cos(t) - D\sin(t). \tag{2.5.8}$$

Substituting $y_p(t)$, $y_p'(t)$, and $y_p''(t)$ into Equation 2.5.4 and simplifying, we find that

$$-2C\sin(t) + 2D\cos(t) = 2\sin(t) \tag{2.5.9}$$

or $D = 0$ and $C = -1$.

To find A and B, we now apply the initial conditions on the general solution

$$y(t) = Ae^{-t} + Bte^{-t} - \cos(t). \tag{2.5.10}$$

The initial condition $y(0) = 2$ yields

$$y(0) = A + 0 - 1 = 2, \tag{2.5.11}$$

or $A = 3$. The initial condition $y'(0) = 1$ gives

$$y'(0) = -A + B = 1, \tag{2.5.12}$$

or $B = 4$, since

$$y'(t) = -Ae^{-t} + Be^{-t} - Bte^{-t} + \sin(t). \tag{2.5.13}$$

Therefore, the solution that satisfies the differential equation and initial conditions is

$$y(t) = 3e^{-t} + 4te^{-t} - \cos(t). \tag{2.5.14}$$

□

• Example 2.5.2

Let us solve the differential equation for a weakly damped harmonic oscillator when the constant forcing F_0 "turns on" at $t = t_0$. The initial conditions are that $x(0) = x_0$ and $x'(0) = v_0$. Mathematically, the problem is

$$x'' + 2\lambda x' + \omega^2 x = \begin{cases} 0, & 0 < t < t_0, \\ F_0, & t_0 < t, \end{cases} \tag{2.5.15}$$

with $x(0) = x_0$ and $x'(0) = v_0$.

To solve Equation 2.5.15, we first divide the time domain into two regions: $0 < t < t_0$ and $t_0 < t$. For $0 < t < t_0$,

$$x(t) = Ae^{-\lambda t}\cos(\omega_d t) + Be^{-\lambda t}\sin(\omega_d t), \tag{2.5.16}$$

where $\omega_d^2 = \omega^2 - \lambda^2$. Upon applying the initial conditions,

$$x(t) = x_0 e^{-\lambda t}\cos(\omega_d t) + \frac{v_0 + \lambda x_0}{\omega_d} e^{-\lambda t}\sin(\omega_d t), \tag{2.5.17}$$

as before.

For the region $t_0 < t$, we write the general solution as

$$\begin{aligned} x(t) = {} & Ae^{-\lambda t}\cos(\omega_d t) + Be^{-\lambda t}\sin(\omega_d t) + \frac{F_0}{\omega^2} \\ & + Ce^{-\lambda(t-t_0)}\cos[\omega_d(t-t_0)] + De^{-\lambda(t-t_0)}\sin[\omega_d(t-t_0)]. \end{aligned} \tag{2.5.18}$$

Why have we written our solution in this particular form rather than the simpler

$$x(t) = Ce^{-\lambda t}\cos(\omega_d t) + De^{-\lambda t}\sin(\omega_d t) + \frac{F_0}{\omega^2}? \tag{2.5.19}$$

Both solutions satisfy the differential equation, as direct substitution verifies. However, the algebra is greatly simplified when Equation 2.5.18 rather than Equation 2.5.19 is used in matching the solution from each region at $t = t_0$. There both the solution and its first derivative must be continuous or

$$x(t_0^-) = x(t_0^+), \qquad \text{and} \qquad x'(t_0^-) = x'(t_0^+), \tag{2.5.20}$$

Review of the Solution of the
Forced Harmonic Oscillator Problem

The undamped system $mx''+kx = F_0\cos(\omega_0 t)$ subject to the initial conditions $x(0) = x_0$ and $x'(0) = v_0$ has the solution

$$x(t) = \frac{v_0}{\omega}\sin(\omega t) + \left(x_0 - \frac{f_0}{\omega^2-\omega_0^2}\right)\cos(\omega t) + \frac{f_0}{\omega^2-\omega_0^2}\cos(\omega_0 t),$$

where $f_0 = F_0/m$ and $\omega = \sqrt{k/m}$. The underdamped system $mx'' + \beta x' + kx = F_0\cos(\omega_0 t)$ has the *steady-state* solution

$$x(t) = \frac{f_0}{\sqrt{(\omega^2-\omega_0^2)^2 + (2\lambda\omega_0)^2}}\cos\left[\omega_0 t - \tan^{-1}\left(\frac{2\lambda\omega_0}{\omega^2-\omega_0^2}\right)\right],$$

where $2\lambda = \beta/m$.

where t_0^- and t_0^+ are points just below and above t_0, respectively. When Equation 2.5.17 and Equation 2.5.18 are substituted, we find that $C = -F_0/\omega^2$, and $\omega_d D = \lambda C$. Thus, the solution for the region $t_0 < t$ is

$$x(t) = x_0 e^{-\lambda t}\cos(\omega_d t) + \frac{v_0+\lambda x_0}{\omega_d}e^{-\lambda t}\sin(\omega_d t) + \frac{F_0}{\omega^2} - \frac{F_0}{\omega^2}e^{-\lambda(t-t_0)}\cos[\omega_d(t-t_0)] - \frac{\lambda F_0}{\omega_d\omega^2}e^{-\lambda(t-t_0)}\sin[\omega_d(t-t_0)]. \quad \textbf{(2.5.21)}$$

As we will see in Chapter 12, the technique of Laplace transforms is particularly well suited for this type of problem when the forcing function changes abruptly at one or more times.□

As noted earlier, nonhomogeneous solutions consist of the homogeneous solution plus a particular solution. In the case of a damped harmonic oscillator, another, more physical, way of describing the solution involves its behavior at large time. That portion of the solution which eventually becomes negligible as $t \to \infty$ is often referred to as the *transient term*, or *transient solution*. In Equation 2.5.14 the transient solution equals $3e^{-t} + 4te^{-t}$. On the other hand, the portion of the solution that remains as $t \to \infty$ is called the *steady-state solution*. In Equation 2.5.14 the steady-state solution equals $-\cos(t)$.

One of the most interesting forced oscillator problems occurs when $\beta = 0$ and the forcing function equals $F_0\sin(\omega_0 t)$, where F_0 is a constant. Then the initial-value problem becomes

$$\frac{d^2x}{dt^2} + \omega^2 x = F_0\sin(\omega_0 t). \quad \textbf{(2.5.22)}$$

Let us solve this problem when $x(0) = x'(0) = 0$.

The homogeneous solution to Equation 2.5.22 is

$$x_H(t) = A\cos(\omega t) + B\sin(\omega t). \quad \textbf{(2.5.23)}$$

To obtain the particular solution, we assume that

$$x_p(t) = C\cos(\omega_0 t) + D\sin(\omega_0 t). \quad \textbf{(2.5.24)}$$

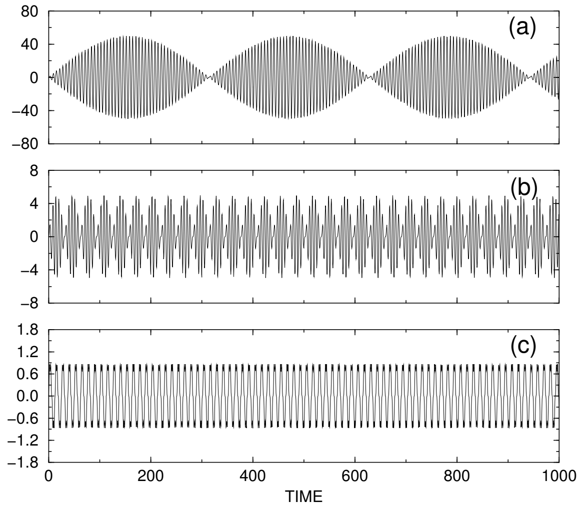


Figure 2.5.1: The solution, Equation 2.5.31, as a function of time when $\omega = 1$ and ω_0 equals (a) 1.02, (b) 1.2, and (c) 2.

This leads to

$$x_p'(t) = -C\omega_0 \sin(\omega_0 t) + D\omega_0 \cos(\omega_0 t), \tag{2.5.25}$$

$$x_p''(t) = -C\omega_0^2 \cos(\omega_0 t) + D\omega_0^2 \sin(\omega_0 t), \tag{2.5.26}$$

and

$$x_p'' + \omega^2 x_p = C(\omega^2 - \omega_0^2)\cos(\omega_0 t) + D(\omega^2 - \omega_0^2)\sin(\omega_0 t) = F_0 \sin(\omega_0 t). \tag{2.5.27}$$

We immediately conclude that $C(\omega^2 - \omega_0^2) = 0$, and $D(\omega^2 - \omega_0^2) = F_0$. Therefore,

$$C = 0, \qquad \text{and} \qquad D = \frac{F_0}{\omega^2 - \omega_0^2}, \tag{2.5.28}$$

provided that $\omega \neq \omega_0$. Thus,

$$x_p(t) = \frac{F_0}{\omega^2 - \omega_0^2} \sin(\omega_0 t). \tag{2.5.29}$$

To finish the problem, we must apply the initial conditions to the general solution

$$x(t) = A\cos(\omega t) + B\sin(\omega t) + \frac{F_0}{\omega^2 - \omega_0^2} \sin(\omega_0 t). \tag{2.5.30}$$

From $x(0) = 0$, we find that $A = 0$. On the other hand, $x'(0) = 0$ yields $B = -\omega_0 F_0/[\omega(\omega^2 - \omega_0^2)]$. Thus, the final result is

$$x(t) = \frac{F_0}{\omega(\omega^2 - \omega_0^2)} \left[\omega \sin(\omega_0 t) - \omega_0 \sin(\omega t)\right]. \tag{2.5.31}$$

Equation 2.5.31 is illustrated in Figure 2.5.1 as a function of time.

The most arresting feature in Figure 2.5.1 is the evolution of the uniform amplitude of the oscillation shown in frame (c) into the one shown in frame (a) where the amplitude exhibits a sinusoidal variation as $\omega_0 \to \omega$. In acoustics these fluctuations in the amplitude are called *beats*, the loud sounds corresponding to the larger amplitudes.

As our analysis indicates, Equation 2.5.31 does not apply when $\omega = \omega_0$. As we shall shortly see, this is probably the most interesting configuration. We can use Equation 2.5.31 to examine this case by applying L'Hôpital's rule in the limiting case of $\omega_0 \to \omega$. This limiting process is analogous to "tuning in" the frequency of the driving frequency $[\omega_0/(2\pi)]$ to the frequency of free vibrations $[\omega/(2\pi)]$. From experience, we expect that given enough time we should be able to substantially increase the amplitudes of vibrations. Mathematical confirmation of our physical intuition is as follows:

$$x(t) = \lim_{\omega_0\to\omega} F_0 \frac{\omega\sin(\omega_0 t) - \omega_0\sin(\omega t)}{\omega(\omega^2-\omega_0^2)} \tag{2.5.32}$$

$$= F_0 \lim_{\omega_0\to\omega} \frac{d[\omega\sin(\omega_0 t) - \omega_0\sin(\omega t)]/d\omega_0}{d[\omega(\omega^2-\omega_0^2)]/d\omega_0} \tag{2.5.33}$$

$$= F_0 \lim_{\omega_0\to\omega} \frac{\omega t\cos(\omega_0 t) - \sin(\omega t)}{-2\omega_0\omega} \tag{2.5.34}$$

$$= F_0 \frac{\omega t\cos(\omega t) - \sin(\omega t)}{-2\omega^2} \tag{2.5.35}$$

$$= \frac{F_0}{2\omega^2}\sin(\omega t) - \frac{F_0 t}{2\omega}\cos(\omega t). \tag{2.5.36}$$

As we suspected, as $t \to \infty$, the displacement grows without bounds. This phenomenon is known as *pure resonance*. We could also have obtained Equation 2.5.36 directly using the method of undetermined coefficients involving the initial value problem

$$\frac{d^2x}{dt^2} + \omega^2 x = F_0\sin(\omega t), \qquad x(0) = x'(0) = 0. \tag{2.5.37}$$

Because there is almost always some friction, pure resonance rarely occurs and the more realistic differential equation is

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + \omega^2 x = F_0\sin(\omega_0 t). \tag{2.5.38}$$

Its solution is

$$x(t) = Ce^{-\lambda t}\sin\left(t\sqrt{\omega^2-\omega_0^2} + \varphi\right) + \frac{F_0}{\sqrt{(\omega^2-\omega_0^2)^2 + 4\lambda^2\omega_0^2}}\sin(\omega_0 t - \theta), \tag{2.5.39}$$

where

$$\sin(\theta) = \frac{2\lambda\omega_0}{\sqrt{(\omega^2-\omega_0^2)^2 + 4\lambda^2\omega_0^2}}, \qquad \cos(\theta) = \frac{\omega^2-\omega_0^2}{\sqrt{(\omega^2-\omega_0^2)^2 + 4\lambda^2\omega_0^2}}, \tag{2.5.40}$$

and C and φ are determined by the initial conditions. To illustrate Equation 2.5.39 we rewrite the amplitude and phase of the particular solution as

$$\frac{F_0}{\sqrt{(\omega^2-\omega_0^2)^2 + 4\lambda^2\omega_0^2}} = \frac{F_0}{\omega^2\sqrt{(1-r^2)^2 + 4\beta^2 r^2}} \qquad \text{and} \qquad \tan(\theta) = \frac{2\beta r}{1-r^2}, \tag{2.5.41}$$

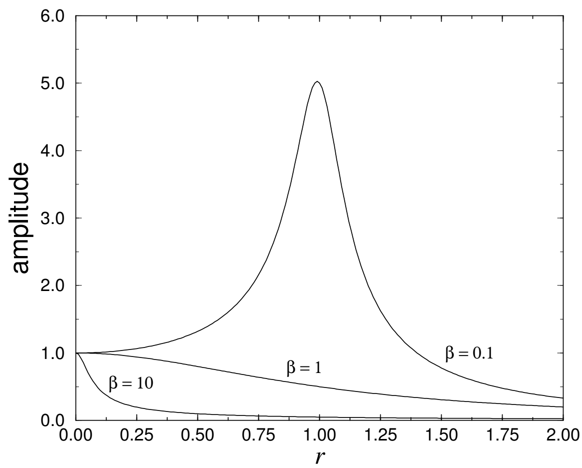


Figure 2.5.2: The amplitude of the particular solution Equation 2.5.39 for a forced, damped simple harmonic oscillator (normalized with F_0/ω^2) as a function of $r = \omega_0/\omega$.

where $r = \omega_0/\omega$ and $\beta = \lambda/\omega$. Figures 2.5.2 and 2.5.3 graph Equation 2.5.41 as functions of r for various values of β.

• Example 2.5.3: Electrical circuits

In the previous chapter, we saw how the mathematical analysis of electrical circuits yields first-order linear differential equations. In those cases we only had a resistor and capacitor or a resistor and inductor. One of the fundamental problems of electrical circuits is a circuit where a resistor, capacitor, and inductor are connected in series, as shown in Figure 2.5.4.

In this RCL circuit, an instantaneous current flows when the key or switch K is closed. If $Q(t)$ denotes the instantaneous charge on the capacitor, Kirchhoff's law yields the differential equation

$$L\frac{dI}{dt} + RI + \frac{Q}{C} = E(t), \qquad \textbf{(2.5.42)}$$

where $E(t)$, the electromotive force, may depend on time, but where L, R, and C are constant. Because $I = dQ/dt$, Equation 2.5.42 becomes

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{Q}{C} = E(t). \qquad \textbf{(2.5.43)}$$

Consider now the case when resistance is negligibly small. Equation 2.5.43 will become identical to the differential equation for the forced simple harmonic oscillator, Equation 2.5.3, with $\lambda = 0$. Similarly, the general case yields various analogs to the damped harmonic oscillator:

Case 1	Overdamped	$R^2 > 4L/C$
Case 2	Critically damped	$R^2 = 4L/C$
Case 3	Underdamped	$R^2 < 4L/C$

In each of these three cases, $Q(t) \to 0$ as $t \to \infty$. (See Problem 6.) Therefore, an RLC electrical circuit behaves like a damped mass-spring mechanical system, where inductance acts like mass, resistance is the damping coefficient, and $1/C$ is the spring constant.

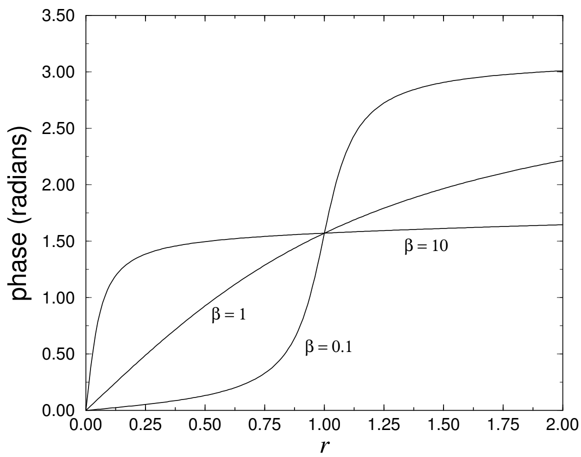


Figure 2.5.3: The phase of the particular solution, Equation 2.5.39, for a forced, damped simple harmonic oscillator as a function of $r = \omega_0/\omega$.

Problems

1. Find the values of γ so that $x'' + 6x' + 18 = \cos(\gamma t)$ is in resonance.

2. The differential equation

$$x'' + 2x' + 2x = 10\sin(2t)$$

describes a damped, forced oscillator. If the initial conditions are $x(0) = x_0$ and $x'(0) = 0$, find its solution by hand and by using MATLAB. Plot the solution when $x_0 = -10, -9, \ldots$,9 , 10. Give a physical interpretation to what you observe.

3. At time $t = 0$, a mass m is suddenly attached to the end of a hanging spring with a spring constant k. Neglecting friction, find the subsequent motion if the coordinate system is chosen so that $x(0) = 0$.

Step 1: Show that the differential equation is

$$m\frac{d^2x}{dt^2} + kx = mg,$$

with the initial conditions $x(0) = x'(0) = 0$.

Step 2: Show that the solution to Step 1 is

$$x(t) = mg\left[1 - \cos(\omega t)\right]/k, \qquad \omega^2 = k/m.$$

4. Consider the electrical circuit shown in Figure 2.5.4, which now possesses negligible resistance and has an applied voltage $E(t) = E_0[1 - \cos(\omega t)]$. Find the *current* if the circuit is initially dead.

5. Find the general solution to the differential equation governing a forced, damped harmonic equation

$$mx'' + cx' + kx = F_0\sin(\omega t),$$

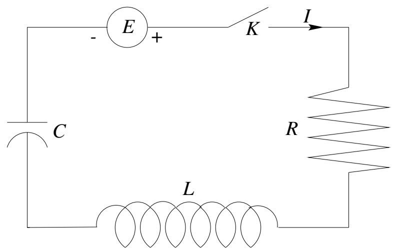


Figure 2.5.4: A simple electrical circuit containing a resistor of constant resistance R, capacitor of constant capacitance C, and inductor of constant inductance L driven by a time-dependent electromotive force $E(t)$.

where m, c, k, F_0, and ω are constants. Write the particular solution in amplitude/phase format.

6. Prove that the *transient* solution to Equation 2.5.45 tends to zero as $t \to \infty$ if R, C, and L are greater than zero.

2.6 VARIATION OF PARAMETERS

As the previous section has shown, the method of undetermined coefficients can be used when the right side of the differential equation contains constants, polynomials, exponentials, sines, and cosines. On the other hand, when the right side contains terms other than these, variation of parameters provides a method for finding the particular solution.

To understand this technique, let us return to our solution of the first-order ordinary differential equation

$$\frac{dy}{dx} + P(x)y = f(x). \tag{2.6.1}$$

Its solution is

$$y(x) = C_1 e^{-\int P(x)\,dx} + e^{-\int P(x)\,dx} \int e^{\int P(x)\,dx} f(x)\,dx. \tag{2.6.2}$$

The solution, Equation 2.6.2, consists of two parts: The first term is the homogeneous solution and can be written $y_H(x) = C_1 y_1(x)$, where $y_1(x) = e^{-\int P(x)\,dx}$. The second term is the particular solution and equals the product of some function of x, say $u_1(x)$, times $y_1(x)$:

$$y_p(x) = e^{-\int P(x)\,dx} \int e^{\int P(x)\,dx} f(x)\,dx = u_1(x) y_1(x). \tag{2.6.3}$$

This particular solution bears a striking resemblance to the homogeneous solution if we replace $u_1(x)$ with C_1.

Variation of parameters builds upon this observation by using the homogeneous solution $y_1(x)$ to construct a guess for the particular solution $y_p(x) = u_1(x)y_1(x)$. Upon substituting this guessed $y_p(x)$ into Equation 2.6.1, we have that

$$\frac{d}{dx}(u_1 y_1) + P(x) u_1 y_1 = f(x), \tag{2.6.4}$$

$$u_1 \frac{dy_1}{dx} + y_1 \frac{du_1}{dx} + P(x) u_1 y_1 = f(x), \tag{2.6.5}$$

or

$$y_1 \frac{du_1}{dx} = f(x), \tag{2.6.6}$$

since $y_1' + P(x)y_1 = 0$.

Using the technique of separating the variables, we have that

$$du_1 = \frac{f(x)}{y_1(x)}\,dx, \qquad \text{and} \qquad u_1(x) = \int \frac{f(x)}{y_1(x)}\,dx. \tag{2.6.7}$$

Consequently, the particular solution equals

$$y_p(x) = u_1(x)y_1(x) = y_1(x)\int \frac{f(x)}{y_1(x)}\,dx. \tag{2.6.8}$$

Upon substituting for $y_1(x)$, we obtain Equation 2.6.3.

How do we apply this method to the linear second-order differential equation

$$a_2(x)y'' + a_1 y'(x) + a_0(x)y = g(x), \tag{2.6.9}$$

or

$$y'' + P(x)y' + Q(x)y = f(x), \tag{2.6.10}$$

where $P(x)$, $Q(x)$, and $f(x)$ are continuous on some interval I?

Let $y_1(x)$ and $y_2(x)$ denote the homogeneous solutions of Equation 2.6.10. That is, $y_1(x)$ and $y_2(x)$ satisfy

$$y_1'' + P(x)y_1' + Q(x)y_1 = 0, \tag{2.6.11}$$

and

$$y_2'' + P(x)y_2' + Q(x)y_2 = 0. \tag{2.6.12}$$

Following our previous example, we now seek two functions $u_1(x)$ and $u_2(x)$ such that

$$y_p(x) = u_1(x)y_1(x) + u_2(x)y_2(x) \tag{2.6.13}$$

is a particular solution of Equation 2.6.10. Once again, we replaced our arbitrary constants C_1 and C_2 by the "variable parameters" $u_1(x)$ and $u_2(x)$. Because we have two unknown functions, we require two equations to solve for $u_1(x)$ and $u_2(x)$. One of them follows from substituting $y_p(x) = u_1(x)y_1(x) + u_2(x)y_2(x)$ into Equation 2.6.10. The other equation is

$$y_1(x)u_1'(x) + y_2(x)u_2'(x) = 0. \tag{2.6.14}$$

This equation is an assumption that is made to simplify the first and second derivative, which is clearly seen by computing

$$y_p' = u_1y_1' + y_1u_1' + u_2y_2' + y_2u_2' = u_1y_1' + u_2y_2', \tag{2.6.15}$$

after applying Equation 2.6.14. Continuing to the second derivative,

$$y_p'' = u_1y_1'' + y_1'u_1' + u_2y_2'' + y_2'u_2'. \tag{2.6.16}$$

Substituting these results into Equation 2.6.10, we obtain

$$\begin{aligned} y_p'' + P(x)y_p' + Q(x)y_p &= u_1y_1'' + y_1'u_1' + u_2y_2'' + y_2'u_2' \\ &\quad + Pu_1y_1' + Pu_2y_2' + Qu_1y_1 + Qu_2y_2, \qquad (2.6.17) \\ &= u_1\left[y_1'' + P(x)y_1' + Q(x)y_1\right] + u_2\left[y_2'' + P(x)y_2' + Q(x)y_2\right] \\ &\qquad + y_1'u_1' + y_2'u_2' = f(x). \qquad (2.6.18) \end{aligned}$$

Hence, $u_1(x)$ and $u_2(x)$ must be functions that also satisfy the condition

$$y_1'u_1' + y_2'u_2' = f(x). \tag{2.6.19}$$

It is important to note that the differential equation must be written so that it conforms to Equation 2.6.10. This may require the division of the differential equation by $a_2(x)$ so that you have the correct $f(x)$.

Equations 2.6.14 and 2.6.19 constitute a linear system of equations for determining the unknown derivatives u_1' and u_2'. By Cramer's rule,[7] the solutions of Equation 2.6.14 and

[7] If you are unfamiliar with Cramer's rule, see Section 3.3.

Equation 2.6.19 equal

$$u_1'(x) = \frac{W_1}{W}, \qquad \text{and} \qquad u_2'(x) = \frac{W_2}{W}, \tag{2.6.20}$$

where

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}, \quad W_1 = \begin{vmatrix} 0 & y_2 \\ f(x) & y_2' \end{vmatrix}, \quad \text{and} \quad W_2 = \begin{vmatrix} y_1 & 0 \\ y_1' & f(x) \end{vmatrix}. \tag{2.6.21}$$

The determinant W is the Wronskian of y_1 and y_2. Because y_1 and y_2 are linearly independent on I, the Wronskian will never equal to zero for every x in the interval.

These results can be generalized to any nonhomogeneous, nth-order, linear equation of the form

$$y^{(n)} + P_{n-1}(x)y^{(n-1)} + P_1(x)y' + P_0(x) = f(x). \tag{2.6.22}$$

If $y_H(x) = C_1y_1(x) + C_2y_2(x) + \cdots + C_ny_n(x)$ is the complementary function for Equation 2.6.22, then a particular solution is

$$y_p(x) = u_1(x)y_1(x) + u_2(x)y_2(x) + \cdots + u_n(x)y_n(x), \tag{2.6.23}$$

where the u_k', $k = 1, 2, \ldots, n$, are determined by the n equations:

$$y_1u_1' + y_2u_2' + \cdots + y_nu_n' = 0, \tag{2.6.24}$$

$$y_1'u_1' + y_2'u_2' + \cdots + y_n'u_n' = 0, \tag{2.6.25}$$

$$\vdots$$

$$y_1^{(n-1)}u_1' + y_2^{(n-1)}u_2' + \cdots + y_n^{(n-1)}u_n' = f(x). \tag{2.6.26}$$

The first $n-1$ equations in this system, like Equation 2.6.14, are assumptions made to simplify the first $n-1$ derivatives of $y_p(x)$. The last equation of the system results from substituting the n derivative of $y_p(x)$ and the simplified lower derivatives into Equation 2.6.22. Then, by Cramer's rule, we find that

$$u_k' = \frac{W_k}{W}, \qquad k = 1, 2, \ldots, n, \tag{2.6.27}$$

where W is the Wronskian of y_1, y_2, ..., y_n, and W_k is the determinant obtained by replacing the kth column of the Wronskian by the column vector $[0, 0, 0, \cdots, f(x)]^T$.

• **Example 2.6.1**

Let us apply variation of parameters to find the general solution to

$$y'' + y' - 2y = xe^x. \tag{2.6.28}$$

We begin by first finding the homogeneous solution that satisfies the differential equation

$$y_H'' + y_H' - 2y_H = 0. \tag{2.6.29}$$

Applying the techniques from Section 2.1, the homogeneous solution is

$$y_H(x) = Ae^x + Be^{-2x}, \tag{2.6.30}$$

yielding the two independent solutions $y_1(x) = e^x$, and $y_2(x) = e^{-2x}$. Thus, the method of variation of parameters yields the particular solution

$$y_p(x) = e^x u_1(x) + e^{-2x} u_2(x). \tag{2.6.31}$$

From Equation 2.6.14, we have that

$$e^x u_1'(x) + e^{-2x} u_2'(x) = 0, \tag{2.6.32}$$

while

$$e^x u_1'(x) - 2e^{-2x} u_2'(x) = xe^x. \tag{2.6.33}$$

Solving for $u_1'(x)$ and $u_2'(x)$, we find that

$$u_1'(x) = \tfrac{1}{3}x, \tag{2.6.34}$$

or

$$u_1(x) = \tfrac{1}{6}x^2, \tag{2.6.35}$$

and

$$u_2'(x) = -\tfrac{1}{3}xe^{3x}, \tag{2.6.36}$$

or

$$u_2(x) = \tfrac{1}{27}(1-3x)e^{3x}. \tag{2.6.37}$$

Therefore, the general solution is

$$\begin{aligned} y(x) &= Ae^x + Be^{-2x} + e^x u_1(x) + e^{-2x} u_2(x) && (2.6.38)\\ &= Ae^x + Be^{-2x} + \tfrac{1}{6}x^2 e^x + \tfrac{1}{27}(1-3x)e^x && (2.6.39)\\ &= Ce^x + Be^{-2x} + \left(\tfrac{1}{6}x^2 - \tfrac{1}{9}x\right)e^x. && (2.6.40)\end{aligned}$$

□

• **Example 2.6.2**

Let us find the general solution to

$$y'' + 2y' + y = e^{-x}\ln(x) \tag{2.6.41}$$

by variation of parameters on the interval $(0, \infty)$.

We start by finding the homogeneous solution that satisfies the differential equation

$$y_H'' + 2y_H' + y_H = 0. \tag{2.6.42}$$

Applying the techniques from Section 2.1, the homogeneous solution is

$$y_H(x) = Ae^{-x} + Bxe^{-x}, \tag{2.6.43}$$

yielding the two independent solutions $y_1(x) = e^{-x}$ and $y_2(x) = xe^{-x}$. Thus, the particular solution equals

$$y_p(x) = e^{-x}u_1(x) + xe^{-x}u_2(x). \tag{2.6.44}$$

From Equation 2.6.14, we have that

$$e^{-x}u_1'(x) + xe^{-x}u_2'(x) = 0, \tag{2.6.45}$$

while

$$-e^{-x}u_1'(x) + (1-x)e^{-x}u_2'(x) = e^{-x}\ln(x). \tag{2.6.46}$$

Solving for $u_1'(x)$ and $u_2'(x)$, we find that

$$u_1'(x) = -x\ln(x), \tag{2.6.47}$$

or

$$u_1(x) = \tfrac{1}{4}x^2 - \tfrac{1}{2}x^2\ln(x), \tag{2.6.48}$$

and

$$u_2'(x) = \ln(x), \tag{2.6.49}$$

or

$$u_2(x) = x\ln(x) - x. \tag{2.6.50}$$

Therefore, the general solution is

$$y(x) = Ae^{-x} + Bxe^{-x} + e^{-x}u_1(x) + xe^{-x}u_2(x) \tag{2.6.51}$$

$$= Ae^{-x} + Bxe^{-x} + \tfrac{1}{2}x^2\ln(x)e^{-x} - \tfrac{3}{4}x^2e^{-x}. \tag{2.6.52}$$

We can verify our result by using the symbolic toolbox in MATLAB. Typing the command:

```
dsolve('D2y+2*Dy+y=exp(-x)*log(x)','x')
```

yields

```
ans =
1/2*exp(-x)*x^2*log(x)-3/4*exp(-x)*x^2+C1*exp(-x)+C2*exp(-x)*x
```
□

• Example 2.6.3

So far, all of our examples have yielded closed-form solutions. To show that this is not necessarily so, let us solve

$$y'' - 4y = e^{2x}/x \tag{2.6.53}$$

by variation of parameters.

Again we begin by solving the homogeneous differential equation

$$y_H'' - 4y_H = 0, \tag{2.6.54}$$

which has the solution

$$y_H(x) = Ae^{2x} + Be^{-2x}. \tag{2.6.55}$$

Thus, our two independent solutions are $y_1(x) = e^{2x}$ and $y_2(x) = e^{-2x}$. Therefore, the particular solution equals

$$y_p(x) = e^{2x}u_1(x) + e^{-2x}u_2(x). \tag{2.6.56}$$

From Equation 2.6.14, we have that

$$e^{2x}u_1'(x) + e^{-2x}u_2'(x) = 0, \tag{2.6.57}$$

while

$$2e^{2x}u_1'(x) - 2e^{-2x}u_2'(x) = e^{2x}/x. \tag{2.6.58}$$

Solving for $u_1'(x)$ and $u_2'(x)$, we find that

$$u_1'(x) = \frac{1}{4x}, \tag{2.6.59}$$

or

$$u_1(x) = \tfrac{1}{4}\ln|x|, \tag{2.6.60}$$

and

$$u_2'(x) = -\frac{e^{4x}}{4x}, \tag{2.6.61}$$

or

$$u_2(x) = -\tfrac{1}{4}\int_{x_0}^{x} \frac{e^{4t}}{t}\,dt. \tag{2.6.62}$$

Therefore, the general solution is

$$y(x) = Ae^{2x} + Be^{-2x} + e^{2x}u_1(x) + e^{-2x}u_2(x) \tag{2.6.63}$$

$$= Ae^{2x} + Be^{-2x} + \tfrac{1}{4}\ln|x|e^{2x} - \tfrac{1}{4}e^{-2x}\int_{x_0}^{x} \frac{e^{4t}}{t}\,dt. \tag{2.6.64}$$

Problems

Use variation of parameters to find the general solution for the following differential equations. Then see if you can obtain your solution by using `dsolve` in MATLAB.

1. $y'' - 4y' + 3y = e^{-x}$
2. $y'' - y' - 2y = x$
3. $y'' - 4y = xe^{x}$
4. $y'' + 9y = 2\sec(x)$
5. $y'' + 4y' + 4y = xe^{-2x}$
6. $y'' + 2ay' = \sin^2(\omega x)$
7. $y'' - 4y' + 4y = (x+1)e^{2x}$
8. $y'' - 4y = \sin^2(x)$
9. $y'' - 2y' + y = e^{x}/x$
10. $y'' + y = \tan(x)$

2.7 EULER-CAUCHY EQUATION

The Euler-Cauchy or equidimensional equation is a linear differential equation of the form

$$a_n x^n \frac{d^n y}{dx^n} + a_{n-1}x^{n-1}\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1 x\frac{dy}{dx} + a_0 y = f(x), \tag{2.7.1}$$

where a_n, a_{n-1}, ..., a_0 are constants. The important point here is that in each term the power to which x is raised equals the *order* of differentiation.

To illustrate this equation, we will focus on the homogeneous, second-order, ordinary differential equation

$$ax^2\frac{d^2y}{dx^2} + bx\frac{dy}{dx} + cy = 0. \tag{2.7.2}$$

The solution of higher-order ordinary differential equations follows by analog. If we wish to solve the nonhomogeneous equation

$$ax^2 \frac{d^2y}{dx^2} + bx\frac{dy}{dx} + cy = f(x), \tag{2.7.3}$$

we can do so by applying variation of parameters using the complementary solutions that satisfy Equation 2.7.2.

Our analysis starts by trying a solution of the form $y = x^m$, where m is presently undetermined. The first and second derivatives are

$$\frac{dy}{dx} = mx^{m-1}, \quad \text{and} \quad \frac{d^2y}{dx^2} = m(m-1)x^{m-2}, \tag{2.7.4}$$

respectively. Consequently, substitution yields the differential equation

$$ax^2 \frac{d^2y}{dx^2} + bx\frac{dy}{dx} + cy = ax^2 \cdot m(m-1)x^{m-2} + bx \cdot mx^{m-1} + cx^m \tag{2.7.5}$$

$$= am(m-1)x^m + bmx^m + cx^m \tag{2.7.6}$$

$$= [am(m-1) + bm + c]\, x^m. \tag{2.7.7}$$

Thus, $y = x^m$ is a solution of the differential equation whenever m is a solution of the *auxiliary equation*

$$am(m-1) + bm + c = 0, \qquad \text{or} \qquad am^2 + (b-a)m + c = 0. \tag{2.7.8}$$

At this point we must consider three different cases that depend upon the values of a, b, and c.

- *Distinct real roots*

Let m_1 and m_2 denote the real roots of Equation 2.7.8 such that $m_1 \neq m_2$. Then,

$$y_1(x) = x^{m_1} \qquad \text{and} \qquad y_2(x) = x^{m_2} \tag{2.7.9}$$

are homogeneous solutions to Equation 2.7.2. Therefore, the general solution is

$$y(x) = C_1 x^{m_1} + C_2 x^{m_2}. \tag{2.7.10}$$

- *Repeated real roots*

If the roots of Equation 2.7.8 are repeated $[m_1 = m_2 = -(b-a)/2]$, then we presently have only one solution, $y = x^{m_1}$. To construct the second solution y_2, we use reduction in order. We begin by first rewriting the Euler-Cauchy equation as

$$\frac{d^2y}{dx^2} + \frac{b}{ax}\frac{dy}{dx} + \frac{c}{ax^2}y = 0. \tag{2.7.11}$$

Letting $P(x) = b/(ax)$, we have

$$y_2(x) = x^{m_1} \int \frac{e^{-\int [b/(ax)]\, dx}}{(x^{m_1})^2}\, dx = x^{m_1} \int \frac{e^{-(b/a)\,\ln(x)}}{x^{2m_1}}\, dx \tag{2.7.12}$$

$$= x^{m_1} \int x^{-b/a} x^{-2m_1}\, dx = x^{m_1} \int x^{-b/a} x^{(b-a)/a}\, dx \tag{2.7.13}$$

$$= x^{m_1} \int \frac{dx}{x} = x^{m_1} \ln(x). \tag{2.7.14}$$

The general solution is then

$$y(x) = C_1 x^{m_1} + C_2 x^{m_1} \ln(x). \tag{2.7.15}$$

For higher-order equations, if m_1 is a root of multiplicity k, then it can be shown that

$$x^{m_1}, x^{m_1}\ln(x), x^{m_1}[\ln(x)]^2, \ldots, x^{m_1}[\ln(x)]^{k-1}$$

are the k linearly independent solutions. Therefore, the general solution of the differential equation equals a linear combination of these k solutions.

- *Conjugate complex roots*

If the roots of Equation 2.7.8 are the complex conjugate pair $m_1 = \alpha + i\beta$, and $m_2 = \alpha - i\beta$, where α and β are real and $\beta > 0$, then a solution is

$$y(x) = C_1 x^{\alpha+i\beta} + C_2 x^{\alpha-i\beta}. \tag{2.7.16}$$

However, because $x^{i\theta} = [e^{\ln(x)}]^{i\theta} = e^{i\theta\ln(x)}$, we have by Euler's formula

$$x^{i\theta} = \cos[\theta\ln(x)] + i\sin[\theta\ln(x)], \tag{2.7.17}$$

and

$$x^{-i\theta} = \cos[\theta\ln(x)] - i\sin[\theta\ln(x)]. \tag{2.7.18}$$

Substitution into Equation 2.7.16 leads to

$$y(x) = C_3 x^{\alpha}\cos[\beta\ln(x)] + C_4 x^{\alpha}\sin[\beta\ln(x)], \tag{2.7.19}$$

where $C_3 = C_1 + C_2$, and $C_4 = iC_1 - iC_2$.

- **Example 2.7.1**

Let us find the general solution to

$$x^2 y'' + 5xy' - 12y = \ln(x) \tag{2.7.20}$$

by the method of undetermined coefficients and variation of parameters.

In the case of undetermined coefficients, we begin by letting $t = \ln(x)$ and $y(x) = Y(t)$. Substituting these variables into Equation 2.7.20, we find that

$$Y'' + 4Y' - 12Y = t. \tag{2.7.21}$$

The homogeneous solution to Equation 2.7.21 is

$$Y_H(t) = A'e^{-6t} + B'e^{2t}, \tag{2.7.22}$$

while the particular solution is

$$Y_p(t) = Ct + D \tag{2.7.23}$$

from the method of undetermined coefficients. Substituting Equation 2.7.23 into Equation 2.7.21 yields $C = -\frac{1}{12}$ and $D = -\frac{1}{36}$. Therefore,

$$Y(t) = A'e^{-6t} + B'e^{2t} - \tfrac{1}{12}t - \tfrac{1}{36}, \tag{2.7.24}$$

or

$$y(x) = \frac{A}{x^6} + Bx^2 - \tfrac{1}{12}\ln(x) - \tfrac{1}{36}. \qquad \textbf{(2.7.25)}$$

To find the particular solution via variation of parameters, we use the homogeneous solution

$$y_H(x) = \frac{A}{x^6} + Bx^2 \qquad \textbf{(2.7.26)}$$

to obtain $y_1(x) = x^{-6}$ and $y_2(x) = x^2$. Therefore,

$$y_p(x) = x^{-6}u_1(x) + x^2u_2(x). \qquad \textbf{(2.7.27)}$$

Substitution of Equation 2.7.27 in Equation 2.7.20 yields the system of equations:

$$x^{-6}u_1'(x) + x^2u_2'(x) = 0, \qquad \textbf{(2.7.28)}$$

and

$$-6x^{-7}u_1'(x) + 2xu_2'(x) = \ln(x)/x^2. \qquad \textbf{(2.7.29)}$$

Solving for $u_1'(x)$ and $u_2'(x)$,

$$u_1'(x) = -\frac{x^5\ln(x)}{8}, \qquad \text{and} \qquad u_2'(x) = -\frac{\ln(x)}{8x^3}. \qquad \textbf{(2.7.30)}$$

The solutions of these equations are

$$u_1(x) = -\frac{x^6\ln(x)}{48} + \frac{x^6}{288}, \qquad \text{and} \qquad u_2(x) = -\frac{\ln(x)}{16x^2} - \frac{1}{32x^2}. \qquad \textbf{(2.7.31)}$$

The general solution then equals

$$y(x) = \frac{A}{x^6} + Bx^2 + x^{-6}u_1(x) + x^2u_2(x) = \frac{A}{x^6} + Bx^2 - \tfrac{1}{12}\ln(x) - \tfrac{1}{36}. \qquad \textbf{(2.7.32)}$$

We can verify this result by using the symbolic toolbox in MATLAB. Typing the command:

```
dsolve('x^2*D2y+5*x*Dy-12*y=log(x)','x')
```

yields

```
ans =
-1/12*log(x)-1/36+C1*x^2+C2/x^6
```

Problems

Find the general solution for the following Euler-Cauchy equations valid over the domain $(-\infty, \infty)$. Then check your answer by using `dsolve` in MATLAB.

1. $x^2y'' + xy' - y = 0$
2. $x^2y'' + 2xy' - 2y = 0$
3. $x^2y'' - 2y = 0$
4. $x^2y'' - xy' + y = 0$
5. $x^2y'' + 3xy' + y = 0$
6. $x^2y'' - 3xy' + 4y = 0$
7. $x^2y'' - y' + 5y = 0$
8. $4x^2y'' + 8xy' + 5y = 0$
9. $x^2y'' + xy' + y = 0$
10. $x^2y'' - 3xy' + 13y = 0$
11. $x^3y''' - 2x^2y'' - 2xy' + 8y = 0$
12. $x^2y'' - 2xy' - 4y = x$

2.8 PHASE DIAGRAMS

In Section 1.6 we showed how solutions to first-order ordinary differential equations could be *qualitatively* solved through the use of the phase line. This concept of qualitatively studying differential equations showed promise as a method for deducing many of the characteristics of the solution to a differential equation without actually solving it. In this section we extend these concepts to second-order ordinary differential equations by introducing the *phase plane*.

Consider the differential equation

$$x'' + \mathrm{sgn}(x) = 0, \tag{2.8.1}$$

where the signum function is defined by Equation 11.2.11. Equation 2.8.1 describes, for example, the motion of an infinitesimal ball rolling in a "V"-shaped trough in a constant gravitational field.[8]

Our analysis begins by introducing the new dependent variable $v = x'$ so that Equation 2.8.1 can be written

$$v\frac{dv}{dx} + \mathrm{sgn}(x) = 0, \tag{2.8.2}$$

since

$$x'' = \frac{d^2x}{dt^2} = \frac{dv}{dt} = \frac{dx}{dt}\frac{dv}{dx} = v\frac{dv}{dx}. \tag{2.8.3}$$

Equation 2.8.2 relates v to x and t has disappeared explicitly from the problem. Integrating Equation 2.8.2 with respect to x, we obtain

$$\int v\,dv + \int \mathrm{sgn}(x)\,dx = C, \tag{2.8.4}$$

or

$$\tfrac{1}{2}v^2 + |x| = C. \tag{2.8.5}$$

Equation 2.8.5 expresses conservation of energy because the first term on the left side of this equation expresses the kinetic energy while the second term gives the potential energy. The value of C depends upon the initial condition $x(0)$ and $v(0)$. Thus, for a specific initial condition, our equation gives the relationship between x and v for the motion corresponding to the initial condition.

Although there is a closed-form solution for Equation 2.8.1, let us imagine that there is none. What could we learn from Equation 2.8.5?

Equation 2.8.5 can be represented in a diagram, called a *phase plane*, where x and v are its axes. A given pair of (x, v) is called a *state* of the system. A given state determines all subsequent states because it serves as initial conditions for any subsequent motion.

For each different value of C, we will obtain a curve, commonly known as *phase paths, trajectories*, or *integral curves*, on the phase plane. In Figure 2.8.1, we used the MATLAB script

```
clear
% set up grid points in the (x,v) plane
[x,v] = meshgrid(-5:0.5:5,-5:0.5:5);
% compute slopes
```

[8] See Lipscomb, T., and R. E. Mickens, 1994: Exact solution to the axisymmetric, constant force oscillator equation. *J. Sound Vib.*, **169**, 138–140.

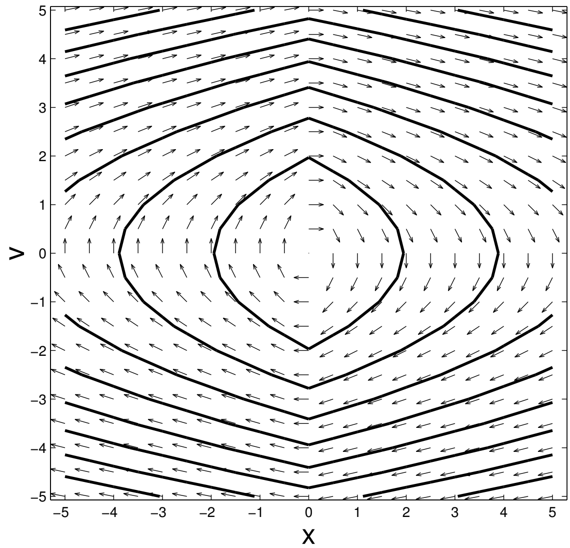


Figure 2.8.1: Phase diagram for the differential equation, Equation 2.8.1.

```
dxdt = v; dvdt = -sign(x);
% find magnitude of vector [dxdt,dydt]
L = sqrt(dxdt.*dxdt + dvdt.*dvdt);
% plot scaled vectors
quiver(x,v,dxdt./L,dvdt./L,0.5); axis equal tight
hold
% contour trajectories
contour(x,v,v.*v/2 + abs(x),8)
h = findobj('Type','patch'); set(h,'Linewidth',2);
xlabel('x','Fontsize',20); ylabel('v','Fontsize',20)
```

to graph the phase plane for Equation 2.8.1. Here the phase paths are simply closed, oval-shaped curves that are symmetric with respect to both the x and v *phase space axes*. Each phase path corresponds to a particular possible motion of the system. Associated with each path is a direction, indicated by an arrow, showing how the state of the system changes as time increases.

An interesting feature on Figure 2.8.1 is the point $(0, 0)$. What is happening there? In our discussion of phase line, we sought to determine whether there were any *equilibrium* or *critical points*. Recall that at an equilibrium or critical point the solution is constant and was given by $x' = 0$. In the case of second-order differential equations, we again have the condition $x' = v = 0$. For this reason, equilibrium points are always situated on the abscissa of the phase diagram.

The condition $x' = 0$ is insufficient for determining critical points. For example, when a ball is thrown upward, its velocity equals zero at the peak height. However, this is clearly not a point of equilibrium. Consequently, we must impose the additional constraint that $x'' = v' = 0$. In the present example, equilibrium points occur where $x' = v = 0$ and $v' = -\text{sgn}(x) = 0$ or $x = 0$. Therefore, the point $(0, 0)$ is the critical point for Equation 2.8.1.

The closed curves immediately surrounding the origin in Figure 2.8.1 show that we have periodic solutions there because on completing a circuit, the original state returns and

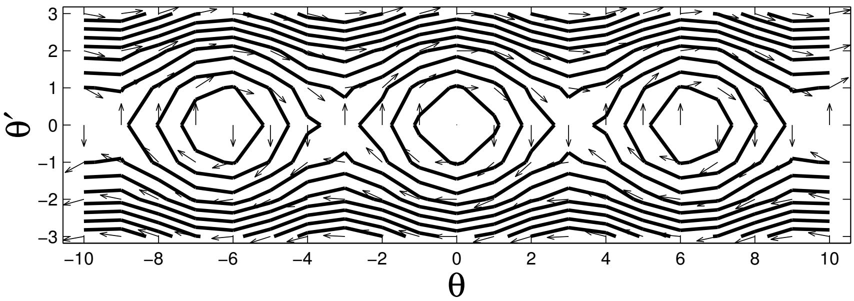


Figure 2.8.2: Phase diagram for a simple pendulum.

the motion simply repeats itself indefinitely.

Once we have found an equilibrium point, an obvious question is whether it is stable or not. To determine this, consider what happens if the initial state is displaced slightly from the origin. It lands on one of the nearby closed curves and the particle oscillates with small amplitude about the origin. Thus, this critical point is *stable*.

In the following examples, we further illustrate the details that may be gleaned from a phase diagram.

• **Example 2.8.1**

The equation describing a simple pendulum is

$$ma^2\theta'' + mga\sin(\theta) = 0, \qquad (\mathbf{2.8.6})$$

where m denotes the mass of the bob, a is the length of the rod or light string, and g is the acceleration due to gravity. Here the conservation of energy equation is

$$\tfrac{1}{2}ma^2\theta'^2 - mga\cos(\theta) = C. \qquad (\mathbf{2.8.7})$$

Figure 2.8.2 is the phase diagram for the simple pendulum. Some of the critical points are located at $\theta = \pm 2n\pi$, $n = 0, 1, 2, \ldots$, and $\theta' = 0$. Near these critical points, we have closed patterns surrounding these critical points, just as we did in the earlier case of an infinitesimal ball rolling in a "V"-shaped trough. Once again, these critical points are *stable* and the region around these equilibrium points corresponds to a pendulum swinging to and fro about the vertical. On the other hand, there is a new type of critical point at $\theta = \pm(2n-1)\pi$, $n = 0, 1, 2, \ldots$ and $\theta' = 0$. Here the trajectories form hyperbolas near these equilibrium points. Thus, for any initial state that is near these critical points, we have solutions that move away from the equilibrium point. This is an example of an *unstable* critical point. Physically these critical points correspond to a pendulum that is balanced on end. Any displacement from the equilibrium results in the bob falling from the inverted position.

Finally, we have a wavy line as $\theta' \to \pm\infty$. This corresponds to whirling motions of the pendulum where θ' has the same sign and θ continuously increases or decreases. □

• **Example 2.8.2: Damped harmonic oscillator**

Consider the ordinary differential equation

$$x'' + 2x' + 5x = 0. \qquad (\mathbf{2.8.8})$$

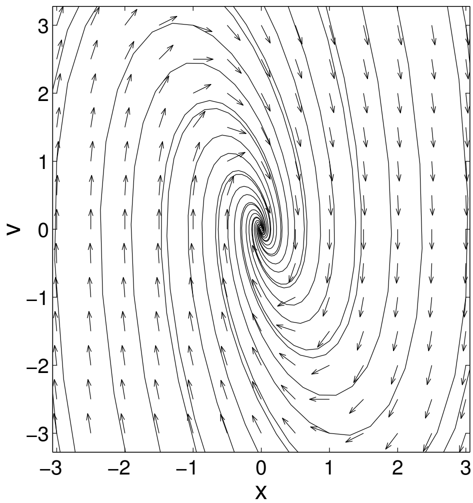


Figure 2.8.3: Phase diagram for the damped harmonic oscillator, Equation 2.8.8.

The exact solution to this differential equation is

$$x(t) = e^{-t}\left[A\cos(2t) + B\sin(2t)\right], \tag{2.8.9}$$

and

$$x'(t) = 2e^{-t}\left[B\cos(2t) - A\sin(2t)\right] - e^{-t}\left[A\cos(2t) + B\sin(2t)\right]. \tag{2.8.10}$$

To construct its phase diagram, we again define $v = x'$ and replace Equation 2.8.8 with $v' = -2v - 5x$. The MATLAB script

```
clear
% set up grid points in the x,x' plane
[x,v] = meshgrid(-3:0.5:3,-3:0.5:3);
% compute slopes
dxdt = v; dvdt = -2*v - 5*x;
% find length of vector
L = sqrt(dxdt.*dxdt + dvdt.*dvdt);
% plot direction field
quiver(x,v,dxdt./L,dvdt./L,0.5); axis equal tight
hold
% compute x(t) and v(t) at various times and a's and b's
for b = -3:2:3; for a = -3:2:3;
  t = [-5:0.1:5];
  xx = exp(-t) .* (a*cos(2*t) + b*sin(2*t));
  vv = 2 * exp(-t) .* (b*cos(2*t) - a*sin(2*t)) - xx;
% plot these values
  plot(xx,vv)
end; end;
xlabel('x','Fontsize',20); ylabel('v','Fontsize',20)
```

was used to construct the phase diagram for Equation 2.8.8 and is shown in Figure 2.8.3. Here the equilibrium point is at $x = v = 0$. This is a new type of critical point. It is called a *stable node* because all slight displacements from this critical point eventually return to this equilibrium point.

Problems

1. Using MATLAB, construct the phase diagram for $x'' - 3x' + 2x = 0$. What happens around the point $x = v = 0$?

2. Consider the nonlinear differential equation $x'' = x^3 - x$. This equation arises in the study of simple pendulums with swings of moderate amplitude.

(a) Show that the conservation law is

$$\tfrac{1}{2}v^2 - \tfrac{1}{4}x^4 + \tfrac{1}{2}x^2 = C.$$

What is special about $C = 0$ and $C = \frac{1}{4}$?

(b) Show that there are three critical points: $x = 0$ and $x = \pm 1$ with $v = 0$.

(c) Using MATLAB, graph the phase diagram with axes x and v.

For the following ordinary differential equations, find the equilibrium points and then classify them. Use MATLAB to draw the phase diagrams.

3. $x'' = 2x'$

4. $x'' + \operatorname{sgn}(x)x = 0$

5. $x'' = \begin{cases} 1, & |x| > 2, \\ 0, & |x| < 2. \end{cases}$

2.9 NUMERICAL METHODS

When differential equations cannot be integrated in closed form, numerical methods must be employed. In the finite difference method, the discrete variable x_i or t_i replaces the continuous variable x or t and the differential equation is solved progressively in increments h starting from known initial conditions. The solution is approximate, but with a sufficiently small increment, you can obtain a solution of acceptable accuracy.

Although there are many different finite difference schemes available, we consider here only two methods that are chosen for their simplicity. The interested student may read any number of texts on numerical analysis if he or she wishes a wider view of other possible schemes.

Let us focus on second-order differential equations; the solution of higher-order differential equations follows by analog. In the case of second-order ordinary differential equations, the differential equation can be rewritten as

$$x'' = f(x, x', t), \qquad x_0 = x(0), \quad x'_0 = x'(0), \tag{2.9.1}$$

where the initial conditions x_0 and x'_0 are assumed to be known.

For the present moment, let us treat the second-order ordinary differential equation

$$x'' = f(x, t), \qquad x_0 = x(0), \quad x'_0 = x'(0). \tag{2.9.2}$$

The following scheme, known as the *central difference method*, computes the solution from Taylor expansions at x_{i+1} and x_{i-1}:

$$x_{i+1} = x_i + hx'_i + \tfrac{1}{2}h^2x''_i + \tfrac{1}{6}h^3x'''_i + O(h^4) \tag{2.9.3}$$

and

$$x_{i-1} = x_i - hx_i' + \tfrac{1}{2}h^2 x_i'' - \tfrac{1}{6}h^3 x_i''' + O(h^4), \qquad (2.9.4)$$

where h denotes the time interval Δt. Subtracting and ignoring higher-order terms, we obtain

$$x_i' = \frac{x_{i+1} - x_{i-1}}{2h}. \qquad (2.9.5)$$

Adding Equation 2.9.3 and Equation 2.9.4 yields

$$x_i'' = \frac{x_{i+1} - 2x_i + x_{i-1}}{h^2}. \qquad (2.9.6)$$

In both Equation 2.9.5 and Equation 2.9.6 we ignored terms of $O(h^2)$. After substituting into the differential equation, Equation 2.9.2, Equation 2.9.6 can be rearranged to

$$x_{i+1} = 2x_i - x_{i-1} + h^2 f(x_i, t_i), \qquad i \ge 1, \qquad (2.9.7)$$

which is known as the *recurrence formula*.

Consider now the situation when $i = 0$. We note that although we have x_0 we do not have x_{-1}. Thus, to start the computation, we need another equation for x_1. This is supplied by Equation 2.9.3, which gives

$$x_1 = x_0 + hx_0' + \tfrac{1}{2}h^2 x_0'' = x_0 + hx_0' + \tfrac{1}{2}h^2 f(x_0, t_0). \qquad (2.9.8)$$

Once we have computed x_1, then we can switch to Equation 2.9.6 for all subsequent calculations.

In this development we have ignored higher-order terms that introduce what is known as *truncation errors*. Other errors, such as *round-off errors*, are introduced due to loss of significant figures. These errors are all related to the time increment h in a rather complicated manner that is investigated in numerical analysis books. In general, better accuracy is obtained by choosing a smaller h, but the number of computations will then increase together with errors.

• Example 2.9.1

Let us solve $x'' - 4x = 2t$ subject to $x(0) = x'(0) = 1$. The exact solution is

$$x(t) = \tfrac{7}{8}e^{2t} + \tfrac{1}{8}e^{-2t} - \tfrac{1}{2}t. \qquad (2.9.9)$$

The MATLAB script

```
clear
% test out different time steps
for i = 1:3
% set up time step increment and number of time steps
  h = 1/10^i; n = 10/h;
% set up initial conditions
  t=zeros(n+1,1); t(1) = 0; x(1) = 1; x_exact(1) = 1;
% define right side of differential equation
  f = inline('4*xx+2*tt','tt','xx');
% set up difference arrays for plotting purposes
  diff = zeros(n,1); t_plot = zeros(n,1);
```

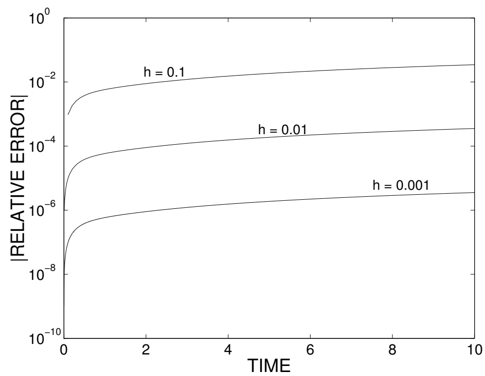


Figure 2.9.1: The numerical solution of $x'' - 4x = 2t$ when $x(0) = x'(0) = 1$ using a simple finite difference approach.

```
% compute first time step
  t(2) = t(1) + h; x(2) = x(1) + h + 0.5*h*h*f(t(1),x(1));
  x_exact(2) = (7/8)*exp(2*t(2))+(1/8)*exp(-2*t(2))-t(2)/2;
  t_plot(1) = t(2);
  diff(1)=x(2)-x_exact(2); diff(1)=abs(diff(1)/x_exact(2));
% compute the remaining time steps
    for k = 2:n
    t(k+1) = t(k) + h; t_plot(k) = t(k+1);
    x(k+1) = 2*x(k) - x(k-1) + h*h*f(t(k),x(k));
    x_exact(k+1) = (7/8)*exp(2*t(k+1))+(1/8)*exp(-2*t(k+1)) ...
                 - t(k+1)/2;
    diff(k) = x(k+1) - x_exact(k+1);
    diff(k) = abs(diff(k) / x_exact(k+1));
  end
% plot the relative error
  semilogy(t_plot,diff,'-')
  hold on
  num = 0.2*n;
  text(3*i,diff(num),['h = ',num2str(h)],'Fontsize',15,...
    'HorizontalAlignment','right','VerticalAlignment','bottom')
  xlabel('TIME','Fontsize',20);
  ylabel('|RELATIVE ERROR|','Fontsize',20);
end
```

implements our simple finite difference method of solving a second-order ordinary differential equation. In Figure 2.9.1 we have plotted results for three different values of the time step. As our analysis suggests, the relative error is related to h^2. □

An alternative method for integrating higher-order ordinary differential equations is Runge-Kutta. It is popular because it is self-starting and the results are very accurate.

For second-order ordinary differential equations, this method first reduces the differential equation into two first-order equations. For example, the differential equation

$$x'' = \frac{f(t) - kx - cx'}{m} = F(x, x', t) \tag{2.9.10}$$

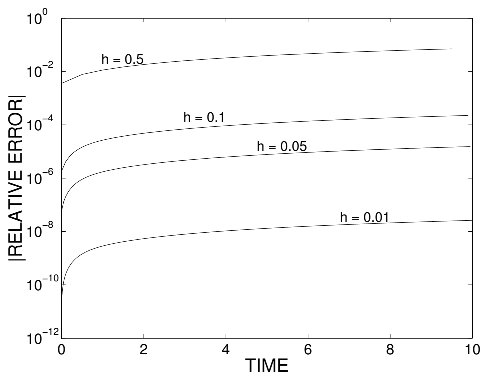


Figure 2.9.2: The numerical solution of $x'' - 4x = 2t$ when $x(0) = x'(0) = 1$ using the Runge-Kutta method.

becomes the first-order differential equations

$$x' = y, \qquad y' = F(x, y, t). \tag{2.9.11}$$

The Runge-Kutta procedure can then be applied to each of these equations. Using a fourth-order scheme, the procedure is as follows:

$$x_{i+1} = x_i + \tfrac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4), \tag{2.9.12}$$

and

$$y_{i+1} = y_i + \tfrac{1}{6}h(K_1 + 2K_2 + 2K_3 + K_4), \tag{2.9.13}$$

where

$$k_1 = y_i, \qquad K_1 = F(x_i, y_i, t_i), \tag{2.9.14}$$

$$k_2 = y_i + \tfrac{h}{2}K_1, \qquad K_2 = F(x_i + \tfrac{h}{2}k_1, k_2, t_i + \tfrac{h}{2}), \tag{2.9.15}$$

$$k_3 = y_i + \tfrac{h}{2}K_2, \qquad K_3 = F(x_i + \tfrac{h}{2}k_2, k_3, t_i + \tfrac{h}{2}), \tag{2.9.16}$$

and

$$k_4 = y_i + K_3 h, \qquad K_4 = F(x_i + hk_3, k_4, t_i + h). \tag{2.9.17}$$

• Example 2.9.2

The MATLAB script

```
clear
% test out different time steps
for i = 1:4
% set up time step increment and number of time steps
  if i==1 h = 0.50; end; if i==2 h = 0.10; end;
  if i==3 h = 0.05; end; if i==4 h = 0.01; end;
  nn = 10/h;
% set up initial conditions
  t=zeros(n+1,1); t(1) = 0;
  x_rk=zeros(n+1,1); x_rk(1) = 1;
  y_rk=zeros(n+1,1); y_rk(1) = 1;
  x_exact=zeros(n+1,1); x_exact(1) = 1;
% set up difference arrays for plotting purposes
  t_plot = zeros(n,1); diff = zeros(n,1);
% define right side of differential equation
  f = inline('4*xx+2*tt','tt','xx','yy');
```

```
  for k = 1:n
    t_local = t(k); x_local = x_rk(k); y_local = y_rk(k);
    k1 = y_local; K1 = f(t_local,x_local,y_local);
    k2 = y_local + h*K1/2;
    K2 = f(t_local + h/2,x_local + h*k1/2,k2);
    k3 = y_local + h*K2/2;
    K3 = f(t_local + h/2,x_local + h*k2/2,k3);
    k4 = y_local + h*K3; K4 = f(t_local + h,x_local + h*k3,k4);
    t(k+1) = t_local + h;
    x_rk(k+1) = x_local + (h/6) * (k1+2*k2+2*k3+k4);
    y_rk(k+1) = y_local + (h/6) * (K1+2*K2+2*K3+K4);
    x_exact(k+1) = (7/8)*exp(2*t(k+1))+(1/8)*exp(-2*t(k+1)) ...
                 - t(k+1)/2;
    t_plot(k) = t(k);
    diff(k) = x_rk(k+1) - x_exact(k+1);
    diff(k) = abs(diff(k) / x_exact(k+1));
  end
% plot the relative errors
  semilogy(t_plot,diff,'-')
  hold on
  xlabel('TIME','Fontsize',20);
  ylabel('|RELATIVE ERROR|','Fontsize',20);
  text(2*i,diff(0.2*n),['h = ',num2str(h)],'Fontsize',15,...
    'HorizontalAlignment','right','VerticalAlignment','bottom')
end
```

was used to resolve Example 2.9.1 using the Runge-Kutta approach. Figure 2.9.2 illustrates the results for time steps of various sizes.

Problems

In previous sections, you found exact solutions to second-order ordinary differential equations. Confirm these earlier results by using MATLAB and the Runge-Kutta scheme to find the numerical solution to the following problems drawn from previous sections.

1. Section 2.1, Problem 1
2. Section 2.1, Problem 5
3. Section 2.4, Problem 1
4. Section 2.4, Problem 5
5. Section 2.6, Problem 1
6. Section 2.6, Problem 5

Project: Pendulum Clock

In his exposition on pendulum clocks, M. Denny[9] modeled the system by the second-order differential equation in time t:

$$\theta'' + b\theta' + \omega_0^2\theta = kf(\theta, \theta'), \tag{1}$$

[9] Denny, M., 2002: The pendulum clock: A venerable dynamical system. *Eur. J. Phys.*, **23**, 449–458.

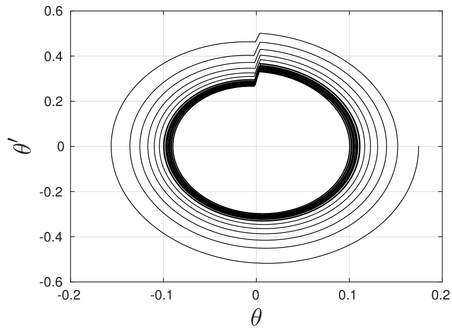


Figure 2.9.3: The phase diagram of the dynamical system given by Equations (1) and (2) by modified Euler method. Here the parameters are $g = 9.8$ m/sec^2, $L = 1$ m, $b = 0.22$/sec, $k = 0.02$/sec, and $\Delta t = 0.005$. The initial conditions are $\theta(0) = \pi/18$ and $\theta'(0) = 0$.

where

$$f(\theta, \theta') = \begin{cases} 1/(\Delta t), & |\theta| < \Delta t/2, \quad \theta' > 0; \\ 0, & \text{otherwise}; \end{cases} \tag{2}$$

and $\omega^2 = g/L - b^2/4$. Here Δt denotes some arbitrarily small nondimensional time. In Chapter 12 we identify this forcing as the Dirac delta function.

Using the numerical scheme of your choice, develop a MATLAB code to numerically integrate this differential equation. Plot the results as a phase diagram with θ as the abscissa and θ' as the ordinate. What happens with time? What happens as k varies? Figure 2.9.3 illustrates the solution.

Further Readings

Boyce, W. E., and R. C. DiPrima, 2004: *Elementary Differential Equations and Boundary Value Problems.* Wiley, 800 pp. Classic textbook.

Ince, E. L., 1956: *Ordinary Differential Equations.* Dover, 558 pp. The source book on ordinary differential equations.

Zill, D. G., and M. R. Cullen, 2008: *Differential Equations with Boundary-Value Problems.* Brooks Cole, 640 pp. Nice undergraduate textbook.

Chapter 3
Linear Algebra

Linear algebra involves the systematic solving of linear algebraic or differential equations that arise during the mathematical modeling of an electrical, mechanical, or even human system where two or more components are interacting with each other. In this chapter we present efficient techniques for expressing these systems and their solutions.

3.1 FUNDAMENTALS OF LINEAR ALGEBRA

Consider the following system of m simultaneous linear equations in n unknowns $x_1, x_2, x_3, \dots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned} \tag{3.1.1}$$

where the coefficients a_{ij} and constants b_j denote known real or complex numbers. The purpose of this chapter is to show how *matrix algebra* can be used to solve these systems by first introducing succinct notation so that we can replace Equation 3.1.1 with rather simple expressions, and then by employing a set of rules to manipulate these expressions. In this section we focus on developing these simple expressions.

The fundamental quantity in linear algebra is the *matrix*.[1] A matrix is an ordered rectangular array of numbers or mathematical expressions. We shall use upper case letters

[1] This term was first used by J. J. Sylvester, 1850: Additions to the articles, "On a new class of theorems," and "On Pascal's theorem." *Philos. Mag., Ser. 4*, **37**, 363–370.

to denote them. The $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_{ij} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & a_{m3} & \cdot & \cdot & \cdot & a_{mn} \end{pmatrix} \tag{3.1.2}$$

has m *rows* and n *columns*. The *order* (or size) of a matrix is determined by the number of rows and columns; Equation 3.1.2 is of order m by n. If $m = n$, the matrix is a *square* matrix; otherwise, A is *rectangular*. The numbers or expressions in the array a_{ij} are the *elements* of A and can be either real or complex. When all of the elements are real, A is a *real matrix*. If some or all of the elements are complex, then A is a *complex matrix*. For a square matrix, the diagonal from the top left corner to the bottom right corner is the *principal diagonal*.

From the limitless number of possible matrices, certain ones appear with sufficient regularity that they are given special names. A *zero* matrix (sometimes called a *null* matrix) has all of its elements equal to zero. It fulfills the role in matrix algebra that is analogous to that of zero in scalar algebra. The *unit* or *identity* matrix is an $n \times n$ matrix having 1's along its principal diagonal and zero everywhere else. The unit matrix serves essentially the same purpose in matrix algebra as does the number one in scalar algebra. A *symmetric* matrix is one where $a_{ij} = a_{ji}$ for all i and j.

• Example 3.1.1

Examples of zero, identity, and symmetric matrices are

$$O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 1 & 0 \\ 4 & 0 & 5 \end{pmatrix}, \tag{3.1.3}$$

respectively. □

A special class of matrices are *column vectors* and *row vectors*:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix}. \tag{3.1.4}$$

We denote row and column vectors by lower case, boldfaced letters. The length or *norm* of the vector $\mathbf{x}$ of n elements is

$$||\mathbf{x}|| = \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2}. \tag{3.1.5}$$

Two matrices A and B are equal if and only if $a_{ij} = b_{ij}$ for all possible i and j and they have the same dimensions.

Having defined a matrix, let us explore some of its arithmetic properties. For two matrices A and B with the same dimensions (conformable for addition), the matrix $C =$

$A + B$ contains the elements $c_{ij} = a_{ij} + b_{ij}$. Similarly, $C = A - B$ contains the elements $c_{ij} = a_{ij} - b_{ij}$. Because the order of addition does not matter, addition is *commutative*: $A + B = B + A$.

Consider now a scalar constant k. The product kA is formed by multiplying every element of A by k. Thus the matrix kA has elements ka_{ij}.

So far the rules for matrix arithmetic conform to their scalar counterparts. However, there are several possible ways of multiplying two matrices together. For example, we might simply multiply together the corresponding elements from each matrix. As we will see, the multiplication rule is designed to facilitate the solution of linear equations.

We begin by requiring that the dimensions of A be $m \times n$ while for B they are $n \times p$. That is, the number of columns in A must equal the number of rows in B. The matrices A and B are then said to be *conformable* for multiplication. If this is true, then $C = AB$ is a matrix $m \times p$, where its elements equal

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}. \tag{3.1.6}$$

The right side of Equation 3.1.6 is referred to as an *inner product* of the ith row of A and the jth column of B. Although Equation 3.1.6 is the method used with a computer, an easier method for human computation is as a running sum of the products given by successive elements of the ith row of A and the corresponding elements of the jth column of B.

The product AA is usually written A^2; the product AAA, A^3, and so forth.

• Example 3.1.2

If

$$A = \begin{pmatrix} -1 & 4 \\ 2 & -3 \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \tag{3.1.7}$$

then

$$AB = \begin{pmatrix} [(-1)(1) + (4)(3)] & [(-1)(2) + (4)(4)] \\ [(2)(1) + (-3)(3)] & [(2)(2) + (-3)(4)] \end{pmatrix} = \begin{pmatrix} 11 & 14 \\ -7 & -8 \end{pmatrix}. \tag{3.1.8}$$

Checking our results using MATLAB, we have that

```
>> A = [-1 4; 2 -3];
>> B = [1 2; 3 4];
>> C = A*B
C =
             11   14
             −7   −8
```

Note that there is a tremendous difference between the MATLAB command for matrix multiplication $*$ and element-by-element multiplication $.*$. □

Matrix multiplication is associative and distributive with respect to addition:

$$(kA)B = k(AB) = A(kB), \tag{3.1.9}$$

$$A(BC) = (AB)C, \tag{3.1.10}$$

$$(A + B)C = AC + BC, \tag{3.1.11}$$

and

$$C(A+B) = CA + CB. \tag{3.1.12}$$

On the other hand, matrix multiplication is *not commutative.* In general, $AB \neq BA$.

• **Example 3.1.3**

Does $AB = BA$ if

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}? \tag{3.1.13}$$

Because

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \tag{3.1.14}$$

and

$$BA = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \tag{3.1.15}$$

$$AB \neq BA. \tag{3.1.16}$$

□

• **Example 3.1.4**

Given

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{3.1.17}$$

find the product AB.

Performing the calculation, we find that

$$AB = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.1.18}$$

The point here is that just because $AB = 0$, this does *not* imply that either A or B equals the zero matrix. □

We cannot properly speak of division when we are dealing with matrices. Nevertheless, a matrix A is said to be *nonsingular* or *invertible* if there exists a matrix B such that $AB = BA = I$. This matrix B is the multiplicative inverse of A or simply the *inverse* of A, written A^{-1}. An $n \times n$ matrix is *singular* if it does not have a multiplicative inverse.

• **Example 3.1.5**

If

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 3 & 4 \\ 2 & 2 & 3 \end{pmatrix}, \tag{3.1.19}$$

let us verify that its inverse is

$$A^{-1} = \begin{pmatrix} 1 & 2 & -3 \\ -1 & 1 & -1 \\ 0 & -2 & 3 \end{pmatrix}. \tag{3.1.20}$$

We perform the check by finding AA^{-1} or $A^{-1}A$,

$$AA^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 3 & 4 \\ 2 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & -3 \\ -1 & 1 & -1 \\ 0 & -2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.1.21}$$

In a later section we will show how to compute the inverse, given A. □

Another matrix operation is transposition. The *transpose* of a matrix A with dimensions $m \times n$ is another matrix, written A^T, where we interchanged the rows and columns from A. In MATLAB, A^T is computed by typing `A'`. Clearly, $(A^T)^T = A$ as well as $(A+B)^T = A^T + B^T$, and $(kA)^T = kA^T$. If A and B are conformable for multiplication, then $(AB)^T = B^T A^T$. Note the reversal of order between the two sides. To prove this last result, we first show that the results are true for two 3×3 matrices A and B and then generalize to larger matrices.

Having introduced some of the basic concepts of linear algebra, we are ready to rewrite Equation 3.1.1 in a canonical form so that we can present techniques for its solution. We begin by writing Equation 3.1.1 as a single column vector:

$$\begin{pmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_m \end{pmatrix}. \tag{3.1.22}$$

We now use the multiplication rule to rewrite Equation 3.1.22 as

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_m \end{pmatrix}, \tag{3.1.23}$$

or

$$A\mathbf{x} = \mathbf{b}, \tag{3.1.24}$$

where $\mathbf{x}$ is the solution vector. If $\mathbf{b} = \mathbf{0}$, we have a *homogeneous* set of equations; otherwise, we have a *nonhomogeneous* set. In the next few sections, we will give a number of methods for finding $\mathbf{x}$.

• Example 3.1.6: Solution of a tridiagonal system

A common problem in linear algebra involves solving systems such as

$$b_1y_1 + c_1y_2 = d_1, \tag{3.1.25}$$

$$a_2y_1 + b_2y_2 + c_2y_3 = d_2, \tag{3.1.26}$$

$$\vdots$$

$$a_{N-1}y_{N-2} + b_{N-1}y_{N-1} + c_{N-1}y_N = d_{N-1}, \tag{3.1.27}$$

$$b_Ny_{N-1} + c_Ny_N = d_N. \tag{3.1.28}$$

Such systems arise in the numerical solution of ordinary and partial differential equations.

We begin by rewriting Equation 3.1.25 through Equation 3.1.28 in the matrix notation:

$$\begin{pmatrix} b_1 & c_1 & 0 & \cdots & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & \cdots & 0 & 0 & 0 \\ 0 & a_3 & b_3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{N-1} & b_{N-1} & c_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & a_N & b_N \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{N-1} \\ y_N \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{N-1} \\ d_N \end{pmatrix}. \tag{3.1.29}$$

The matrix in Equation 3.1.29 is an example of a *banded matrix*: a matrix where all of the elements in each row are zero except for the diagonal element and a limited number on either side of it. In the present case, we have a *tridiagonal* matrix in which only the diagonal element and the elements immediately to its left and right in each row are nonzero.

Consider the nth equation. We can eliminate a_n by multiplying the $(n-1)$th equation by a_n/b_{n-1} and subtracting this new equation from the nth equation. The values of b_n and d_n become

$$b'_n = b_n - a_n c_{n-1}/b_{n-1}, \qquad \text{and} \qquad d'_n = d_n - a_n d_{n-1}/b_{n-1} \tag{3.1.30}$$

for $n = 2, 3, \ldots, N$. The coefficient c_n is unaffected. Because elements a_1 and c_N are never involved, their values can be anything or they can be left undefined. The new system of equations may be written

$$\begin{pmatrix} b'_1 & c_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & b'_2 & c_2 & \cdots & 0 & 0 & 0 \\ 0 & 0 & b'_3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & b'_{N-1} & c_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & b'_N \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{N-1} \\ y_N \end{pmatrix} = \begin{pmatrix} d'_1 \\ d'_2 \\ d'_3 \\ \vdots \\ d'_{N-1} \\ d'_N \end{pmatrix}. \tag{3.1.31}$$

The matrix in Equation 3.1.31 is in *upper triangular* form because all of the elements below the principal diagonal are zero. This is particularly useful because y_n can be computed by *back substitution*. That is, we first compute y_N. Next, we calculate y_{N-1} in terms of y_N. The solution y_{N-2} can then be computed in terms of y_N and y_{N-1}. We continue this process until we find y_1 in terms of $y_N, y_{N-1}, \ldots, y_2$. In the present case, we have the rather simple:

$$y_N = d'_N/b'_N, \qquad \text{and} \qquad y_n = (d'_n - c_n d'_{n+1})/b'_n \tag{3.1.32}$$

for $n = N-1, N-2, \ldots, 2, 1$.

As we shall show shortly, this is an example of solving a system of linear equations by Gaussian elimination. For a tridiagonal case, we have the advantage that the solution can be expressed in terms of a recurrence relationship, a very convenient feature from a computational point of view. This algorithm is very robust, being stable[2] as long as $|a_i + c_i| < |b_i|$. By stability, we mean that if we change $\mathbf{b}$ by $\Delta\mathbf{b}$ so that $\mathbf{x}$ changes by $\Delta\mathbf{x}$, then $||\Delta\mathbf{x}|| < M\epsilon$, where $||\Delta\mathbf{b}|| \le \epsilon$, $0 < M < \infty$, for *any* N.

---

[2] Torii, T., 1966: Inversion of tridiagonal matrices and the stability of tridiagonal systems of linear systems. *Tech. Rep. Osaka Univ.*, **16**, 403–414.

• **Example 3.1.7: Linear transformation**

Consider a set of linear equations

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{14}x_4, \\ y_3 &= a_{31}x_1 + a_{31}x_2 + a_{33}x_3 + a_{34}x_4. \end{aligned} \qquad (\mathbf{3.1.33})$$

Each of the right-side expressions is called a *linear combination* of x_1, x_2, x_3, and x_4: a sum where each term consists of a constant times x_i raised to the first power. An expression such as $a_{11}x_1^2 + a_{12}x_2^2 + a_{13}x_3^2 + a_{14}x_4^2$ is an example of a nonlinear combination. Note that we are using 4 values of x_i to find only 3 values of y_i.

If we were given values of x_i, we can determine a set of values for y_i using Equation 3.1.33. Such a set of linear equations that yields values of y_i for given x_i's is called a *linear transform of* x *into* y. The point here is that given x, the corresponding y will be evaluated.

Matrix notation and multiplication are very convenient in expressing linear transformation. In the present case, we would have that

$$\mathbf{y} = A\mathbf{x}, \quad \text{where} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}. \qquad (\mathbf{3.1.34})$$

In general, a transformation $A(\mathbf{x})$ is a linear transformation that satisfies two conditions: (1) $A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y})$ and (2) $A(k\mathbf{x}) = kA(\mathbf{x})$, where k is a scalar.

Problems

Given $A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$, and $B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, find

1. $A + B$, $B + A$
2. $A - B$, $B - A$
3. $3A - 2B$, $3(2A - B)$
4. A^T,B^T,$(B^T)^T$
5. $(A + B)^T$, $A^T + B^T$
6. $B + B^T$, $B - B^T$
7. AB,A^TB,BA,B^TA
8. A^2, B^2
9. BB^T, B^TB
10. $A^2 - 3A + I$
11. $A^3 + 2A$
12. $A^4 - 4A^2 + 2I$

by hand and using MATLAB.

Can multiplication occur between the following matrices? If so, compute it.

13. $\begin{pmatrix} 3 & 5 & 1 \\ -2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 4 & 1 \\ 1 & 3 \end{pmatrix}$
14. $\begin{pmatrix} -2 & 4 \\ -4 & 6 \\ -6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$
15. $\begin{pmatrix} 1 & 4 & 2 \\ 0 & 0 & 4 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$
16. $\begin{pmatrix} 4 & 6 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 3 & 6 \\ 1 & 2 & 5 \end{pmatrix}$
17. $\begin{pmatrix} 6 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 6 \end{pmatrix}$

If $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 3 & 1 \end{pmatrix}$, verify that

18. $7A = 4A + 3A$, 19. $10A = 5(2A)$, 20. $(A^T)^T = A$
by hand and using MATLAB.

If $A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -2 \\ 4 & 0 \end{pmatrix}$, and $C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, verify that

21. $(A + B) + C = A + (B + C)$, 22. $(AB)C = A(BC)$,

23. $A(B + C) = AB + AC$, 24. $(A + B)C = AC + BC$
by hand and using MATLAB.

Verify that the following A^{-1} are indeed the inverse of A:

25. $A = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$ $A^{-1} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$ 26. $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $A^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
by hand and using MATLAB.

Write the following linear systems of equations in matrix form: $A\mathbf{x} = \mathbf{b}$.

27. $x_1 - 2x_2 = 5$ $3x_1 + x_2 = 1$

28. $2x_1 + x_2 + 4x_3 = 2$ $4x_1 + 2x_2 + 5x_3 = 6$ $6x_1 - 3x_2 + 5x_3 = 2$

29. $x_2 + 2x_3 + 3x_4 = 2$ $3x_1 - 4x_3 - 4x_4 = 5$ $x_1 + x_2 + x_3 + x_4 = -3$
$2x_1 - 3x_2 + x_3 - 3x_4 = 7$

3.2 DETERMINANTS

Determinants appear naturally during the solution of simultaneous equations. Consider, for example, two simultaneous equations with two unknowns x_1 and x_2,

$$a_{11}x_1 + a_{12}x_2 = b_1, \tag{3.2.1}$$

and

$$a_{21}x_1 + a_{22}x_2 = b_2. \tag{3.2.2}$$

The solution to these equations for the value of x_1 and x_2 is

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \tag{3.2.3}$$

and

$$x_2 = \frac{b_2a_{11} - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}. \tag{3.2.4}$$

Note that the denominator of Equation 3.2.3 and Equation 3.2.4 is the same. This term, which always appears in the solution of 2×2 systems, is formally given the name of *determinant* and written

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{3.2.5}$$

MATLAB provides a simple command `det(A)`, which computes the determinant of `A`. For example, in the present case,

```
>> A = [2 -1 2; 1 3 2; 5 1 6];
>> det(A)
ans =
     0
```

Although determinants have their origin in the solution of systems of equations, any square array of numbers or expressions possesses a unique determinant, independent of whether it is involved in a system of equations or not. This determinant is evaluated (or expanded) according to a formal rule known as *Laplace's expansion of cofactors*.[3] The process revolves around expanding the determinant using any arbitrary column *or* row of A. If the ith row or jth column is chosen, the determinant is given by

$$\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}, \quad (\mathbf{3.2.6})$$

where A_{ij}, the *cofactor* of a_{ij}, equals $(-1)^{i+j}M_{ij}$. The minor M_{ij} is the determinant of the $(n-1)\times(n-1)$ submatrix obtained by deleting row i, column j of A. This rule, of course, was chosen so that determinants are still useful in solving systems of equations.

• Example 3.2.1

Let us evaluate

$$\begin{vmatrix} 2 & -1 & 2 \\ 1 & 3 & 2 \\ 5 & 1 & 6 \end{vmatrix}$$

by an expansion in cofactors.

Using the first column,

$$\begin{vmatrix} 2 & -1 & 2 \\ 1 & 3 & 2 \\ 5 & 1 & 6 \end{vmatrix} = 2(-1)^2 \begin{vmatrix} 3 & 2 \\ 1 & 6 \end{vmatrix} + 1(-1)^3 \begin{vmatrix} -1 & 2 \\ 1 & 6 \end{vmatrix} + 5(-1)^4 \begin{vmatrix} -1 & 2 \\ 3 & 2 \end{vmatrix} \quad (\mathbf{3.2.7})$$

$$= 2(16) - 1(-8) + 5(-8) = 0. \quad (\mathbf{3.2.8})$$

The greatest source of error is forgetting to take the factor $(-1)^{i+j}$ into account during the expansion. □

Although Laplace's expansion does provide a method for calculating $\det(A)$, the number of calculations equals $n!$. Consequently, for hand calculations, an obvious strategy is to select the column or row that has the greatest number of zeros. An even better strategy would be to manipulate a determinant with the goal of introducing zeros into a particular column or row. In the remaining portion of this section, we show some operations that may be performed on a determinant to introduce the desired zeros. Most of the properties follow from the expansion of determinants by cofactors.

• *Rule 1* : For every square matrix A, $\det(A^T) = \det(A)$.

---

[3] Laplace, P. S., 1772: Recherches sur le calcul intégral et sur le système du monde. *Hist. Acad. R. Sci., II^e Partie*, 267–376. *Œuvres*, **8**, pp. 369–501. See Muir, T., 1960: *The Theory of Determinants in the Historical Order of Development, Vol. I, Part 1, General Determinants Up to 1841*. Dover Publishers, pp. 24–33.

The proof is left as an exercise.

• *Rule 2* : If any two rows or columns of A are identical, $\det(A) = 0$.

To see that this is true, consider the following 3×3 matrix:

$$\begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix} = c_1(b_2b_3 - b_3b_2) - c_2(b_1b_3 - b_3b_1) + c_3(b_1b_2 - b_2b_1) = 0. \quad \textbf{(3.2.9)}$$

• *Rule 3* : The determinant of a triangular matrix is equal to the product of its diagonal elements.

If A is lower triangular, successive expansions by elements in the first column give

$$\det(A) = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = \cdots = a_{11}a_{22}\cdots a_{nn}. \quad \textbf{(3.2.10)}$$

If A is upper triangular, successive expansions by elements of the first row prove the property.

• *Rule 4* : If a square matrix A has either a row or a column of all zeros, then $\det(A) = 0$.

The proof is left as an exercise.

• *Rule 5* : If each element in one row (column) of a determinant is multiplied by a number c, the value of the determinant is multiplied by c.

Suppose that $|B|$ has been obtained from $|A|$ by multiplying row i (column j) of $|A|$ by c. Upon expanding $|B|$ in terms of row i (column j), each term in the expansion contains c as a factor. Factor out the common c, and the result is just c times the expansion $|A|$ by the same row (column).

• *Rule 6* : If each element of a row (or a column) of a determinant can be expressed as a binomial, the determinant can be written as the sum of two determinants.

To understand this property, consider the following 3×3 determinant:

$$\begin{vmatrix} a_1 + d_1 & b_1 & c_1 \\ a_2 + d_2 & b_2 & c_2 \\ a_3 + d_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}. \quad \textbf{(3.2.11)}$$

The proof follows by expanding the determinant by the row (or column) that contains the binomials.

• *Rule 7* : If B is a matrix obtained by interchanging any two rows (columns) of a square matrix A, then $\det(B) = -\det(A)$.

The proof is by induction. It is easily shown for any 2×2 matrix. Assume that this rule holds for any $(n-1) \times (n-1)$ matrix. If A is $n \times n$, then let B be a matrix formed by interchanging rows i and j. Expanding $|B|$ and $|A|$ by a different row, say k, we have that

$$|B| = \sum_{s=1}^{n} (-1)^{k+s} b_{ks} M_{ks}, \quad \text{and} \quad |A| = \sum_{s=1}^{n} (-1)^{k+s} a_{ks} N_{ks}, \qquad (\mathbf{3.2.12})$$

where M_{ks} and N_{ks} are the minors formed by deleting row k, column s from $|B|$ and $|A|$, respectively. For $s = 1, 2, \ldots, n$, we obtain N_{ks} and M_{ks} by interchanging rows i and j. By the induction hypothesis and recalling that N_{ks} and M_{ks} are $(n-1) \times (n-1)$ determinants, $N_{ks} = -M_{ks}$ for $s = 1, 2, \ldots, n$. Hence, $|B| = -|A|$. Similar arguments hold if two columns are interchanged.

• *Rule 8* : If one row (column) of a square matrix A equals to a number c times some other row (column), then $\det(A) = 0$.

Suppose one row of a square matrix A is equal to c times some other row. If $c = 0$, then $|A| = 0$. If $c \neq 0$, then $|A| = c|B|$, where $|B| = 0$ because $|B|$ has two identical rows. A similar argument holds for two columns.

• *Rule 9* : The value of $\det(A)$ is unchanged if any arbitrary multiple of any line (row or column) is added to any other line.

To see that this is true, consider the simple example:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} cb_1 & b_1 & c_1 \\ cb_2 & b_2 & c_2 \\ cb_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 + cb_1 & b_1 & c_1 \\ a_2 + cb_2 & b_2 & c_2 \\ a_3 + cb_3 & b_3 & c_3 \end{vmatrix}, \qquad (\mathbf{3.2.13})$$

where $c \neq 0$. The first determinant on the left side is our original determinant. In the second determinant, we again expand the first column and find that

$$\begin{vmatrix} cb_1 & b_1 & c_1 \\ cb_2 & b_2 & c_2 \\ cb_3 & b_3 & c_3 \end{vmatrix} = c \begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix} = 0. \qquad (\mathbf{3.2.14})$$

• **Example 3.2.2**

Let us evaluate

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & 1 & 2 & 3 \\ 1 & -1 & 1 & 2 \\ -1 & 1 & -1 & 5 \end{vmatrix}$$

using a combination of the properties stated above and expansion by cofactors.

By adding or subtracting the first row to the other rows, we have that

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & 1 & 2 & 3 \\ 1 & -1 & 1 & 2 \\ -1 & 1 & -1 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 5 & 7 \\ 0 & -3 & -2 & -2 \\ 0 & 3 & 2 & 9 \end{vmatrix} = \begin{vmatrix} 3 & 5 & 7 \\ -3 & -2 & -2 \\ 3 & 2 & 9 \end{vmatrix} \qquad (\mathbf{3.2.15})$$

$$= \begin{vmatrix} 3 & 5 & 7 \\ 0 & 3 & 5 \\ 0 & -3 & 2 \end{vmatrix} = 3 \begin{vmatrix} 3 & 5 \\ -3 & 2 \end{vmatrix} = 3 \begin{vmatrix} 3 & 5 \\ 0 & 7 \end{vmatrix} = 63. \qquad (\mathbf{3.2.16})$$

Problems

Evaluate the following determinants. Check your answer using MATLAB.

1. $\begin{vmatrix} 3 & 5 \\ -2 & -1 \end{vmatrix}$ 2. $\begin{vmatrix} 5 & -1 \\ -8 & 4 \end{vmatrix}$ 3. $\begin{vmatrix} 3 & 1 & 2 \\ 2 & 4 & 5 \\ 1 & 4 & 5 \end{vmatrix}$ 4. $\begin{vmatrix} 4 & 3 & 0 \\ 3 & 2 & 2 \\ 5 & -2 & -4 \end{vmatrix}$

5. $\begin{vmatrix} 1 & 3 & 2 \\ 4 & 1 & 1 \\ 2 & 1 & 3 \end{vmatrix}$ 6. $\begin{vmatrix} 2 & -1 & 2 \\ 1 & 3 & 3 \\ 5 & 1 & 6 \end{vmatrix}$ 7. $\begin{vmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 6 & 1 & 0 \\ 1 & 1 & -2 & 3 \end{vmatrix}$ 8. $\begin{vmatrix} 2 & 1 & 2 & 1 \\ 3 & 0 & 2 & 2 \\ -1 & 2 & -1 & 1 \\ -3 & 2 & 3 & 1 \end{vmatrix}$

9. Using the properties of determinants, show that

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix} = (b-a)(c-a)(d-a)(c-b)(d-b)(d-c).$$

This determinant is called *Vandermonde's determinant*.

10. Show that

$$\begin{vmatrix} a & b+c & 1 \\ b & a+c & 1 \\ c & a+b & 1 \end{vmatrix} = 0.$$

11. Show that if all of the elements of a row or column are zero, then $\det(A) = 0$.

12. Prove that $\det(A^T) = \det(A)$.

3.3 CRAMER'S RULE

One of the most popular methods for solving simple systems of linear equations is Cramer's rule.[4] It is very useful for 2×2 systems, acceptable for 3×3 systems, and of doubtful use for 4×4 or larger systems.

Let us have n equations with n unknowns, $A\mathbf{x} = \mathbf{b}$. Cramer's rule states that

$$x_1 = \frac{\det(A_1)}{\det(A)}, \quad x_2 = \frac{\det(A_2)}{\det(A)}, \quad \cdots, \quad x_n = \frac{\det(A_n)}{\det(A)}, \tag{3.3.1}$$

where A_i is a matrix obtained from A by replacing the ith column with $\mathbf{b}$ and n is the number of unknowns and equations. Obviously, $\det(A) \neq 0$ if Cramer's rule is to work.

[4] Cramer, G., 1750: *Introduction à l'analyse des lignes courbes algébriques*. Geneva, p. 657.

To prove[5] Cramer's rule, consider

$$x_1 \det(A) = \begin{vmatrix} a_{11}x_1 & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21}x_1 & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31}x_1 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}x_1 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \quad (\mathbf{3.3.2})$$

by Rule 5 from the previous section. By adding x_2 times the second column to the first column,

$$x_1 \det(A) = \begin{vmatrix} a_{11}x_1 + a_{12}x_2 & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21}x_1 + a_{22}x_2 & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31}x_1 + a_{32}x_2 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}. \quad (\mathbf{3.3.3})$$

Multiplying each of the columns by the corresponding x_i and adding it to the first column yields

$$x_1 \det(A) = \begin{vmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}. \quad (\mathbf{3.3.4})$$

The first column of Equation 3.3.4 equals $A\mathbf{x}$ and we replace it with $\mathbf{b}$. Thus,

$$x_1 \det(A) = \begin{vmatrix} b_1 & a_{12} & a_{13} & \cdots & a_{1n} \\ b_2 & a_{22} & a_{23} & \cdots & a_{2n} \\ b_3 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_n & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = \det(A_1), \quad (\mathbf{3.3.5})$$

or

$$x_1 = \frac{\det(A_1)}{\det(A)} \quad (\mathbf{3.3.6})$$

provided $\det(A) \neq 0$. To complete the proof we do exactly the same procedure to the jth column.

• Example 3.3.1

Let us solve the following system of equations by Cramer's rule:

$$2x_1 + x_2 + 2x_3 = -1, \quad (\mathbf{3.3.7})$$

$$x_1 + x_3 = -1, \quad (\mathbf{3.3.8})$$

[5] First proved by Cauchy, L. A., 1815: Mémoire sur les fonctions quine peuvent obtemir que deux valeurs égales et de signes contraires par suite des transportations opérées entre les variables qúelles renferment. *J. l'École Polytech.*, **10**, 29–112.

and

$$-x_1 + 3x_2 - 2x_3 = 7. \tag{3.3.9}$$

From the matrix form of the equations,

$$\begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ -1 & 3 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 7 \end{pmatrix}, \tag{3.3.10}$$

we have that

$$\det(A) = \begin{vmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ -1 & 3 & -2 \end{vmatrix} = 1, \tag{3.3.11}$$

$$\det(A_1) = \begin{vmatrix} -1 & 1 & 2 \\ -1 & 0 & 1 \\ 7 & 3 & -2 \end{vmatrix} = 2, \tag{3.3.12}$$

$$\det(A_2) = \begin{vmatrix} 2 & -1 & 2 \\ 1 & -1 & 1 \\ -1 & 7 & -2 \end{vmatrix} = 1, \tag{3.3.13}$$

and

$$\det(A_3) = \begin{vmatrix} 2 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & 3 & 7 \end{vmatrix} = -3. \tag{3.3.14}$$

Finally,

$$x_1 = \frac{2}{1} = 2, \quad x_2 = \frac{1}{1} = 1, \quad \text{and} \quad x_3 = \frac{-3}{1} = -3. \tag{3.3.15}$$

You can also use MATLAB to perform Cramer's rule. In the present example, the script is as follows:

```
clear; % clear all previous computations
A = [2 1 2; 1 0 1; -1 3 -2]; % input coefficient matrix
b = [-1 ; -1; 7]; % input right side
A1 = A; A1(:,1) = b; % compute A_1
A2 = A; A2(:,2) = b; % compute A_2
A3 = A; A3(:,3) = b; % compute A_3
% compute solution vector
x = [det(A1), det(A2), det(A3)] / det(A)
```

Problems

Solve the following systems of equations by Cramer's rule:

1. $x_1 + 2x_2 = 3, \qquad 3x_1 + x_2 = 6$
2. $2x_1 + x_2 = -3, \qquad x_1 - x_2 = 1$
3. $x_1 + 2x_2 - 2x_3 = 4, \qquad 2x_1 + x_2 + x_3 = -2, \qquad -x_1 + x_2 - x_3 = 2$
4. $2x_1 + 3x_2 - x_3 = -1, \qquad -x_1 - 2x_2 + x_3 = 5, \qquad 3x_1 - x_2 = -2.$

Check your answer using MATLAB.

3.4 ROW ECHELON FORM AND GAUSSIAN ELIMINATION

So far, we assumed that every system of equations has a unique solution. This is not necessarily true, as the following examples show.

• Example 3.4.1

Consider the system

$$x_1 + x_2 = 2 \qquad (\mathbf{3.4.1})$$

and

$$2x_1 + 2x_2 = -1. \qquad (\mathbf{3.4.2})$$

This system is inconsistent because the second equation does not follow after multiplying the first by 2. Geometrically, Equation 3.4.1 and Equation 3.4.2 are parallel lines; they never intersect to give a unique x_1 and x_2. □

• Example 3.4.2

Even if a system is consistent, it still may not have a unique solution. For example, the system

$$x_1 + x_2 = 2 \qquad (\mathbf{3.4.3})$$

and

$$2x_1 + 2x_2 = 4 \qquad (\mathbf{3.4.4})$$

is consistent, with the second equation formed by multiplying the first by 2. However, there are an infinite number of solutions. □

Our examples suggest the following:

Theorem: *A system of m linear equations in n unknowns may: (1) have no solution, in which case it is called an* inconsistent *system, or (2) have exactly one solution (called a unique solution), or (3) have an infinite number of solutions. In the latter two cases, the system is said to be* consistent.

Before we can prove this theorem at the end of this section, we need to introduce some new concepts.

The first one is equivalent systems. Two systems of equations involving the same variables are *equivalent* if they have the same solution set. Of course, the only reason for introducing equivalent systems is the possibility of transforming one system of linear systems into another that is easier to solve. But what operations are permissible? Also, what is the ultimate goal of our transformation?

From a complete study of possible operations, there are only three operations for transforming one system of linear equations into another. These three *elementary row operations* are

(1) interchanging any two rows in the matrix,

(2) multiplying any row by a nonzero scalar, and

(3) adding any arbitrary multiple of any row to any other row.

Armed with our elementary row operations, let us now solve the following set of linear equations:

$$x_1 - 3x_2 + 7x_3 = 2, \tag{3.4.5}$$

$$2x_1 + 4x_2 - 3x_3 = -1, \tag{3.4.6}$$

and

$$-x_1 + 13x_2 - 21x_3 = 2. \tag{3.4.7}$$

We begin by writing Equation 3.4.5 through Equation 3.4.7 in matrix notation:

$$\begin{pmatrix} 1 & -3 & 7 \\ 2 & 4 & -3 \\ -1 & 13 & -21 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}. \tag{3.4.8}$$

The matrix in Equation 3.4.8 is called the *coefficient matrix* of the system.

We now introduce the concept of the *augmented matrix*: a matrix B composed of A plus the column vector $\mathbf{b}$ or

$$B = \left(\begin{array}{ccc|c} 1 & -3 & 7 & 2 \\ 2 & 4 & -3 & -1 \\ -1 & 13 & -21 & 2 \end{array}\right). \tag{3.4.9}$$

We can solve our original system by performing elementary row operations on the augmented matrix. Because x_i functions essentially as a placeholder, we can omit them until the end of the computation.

Returning to the problem, the first row can be used to eliminate the elements in the first column of the remaining rows. For this reason the first row is called the *pivotal* row and the element a_{11} is the *pivot*. By using the third elementary row operation twice (to eliminate the 2 and -1 in the first column), we have the equivalent system

$$B = \left(\begin{array}{ccc|c} 1 & -3 & 7 & 2 \\ 0 & 10 & -17 & -5 \\ 0 & 10 & -14 & 4 \end{array}\right). \tag{3.4.10}$$

At this point we choose the second row as our new pivotal row and again apply the third row operation to eliminate the last element in the second column. This yields

$$B = \left(\begin{array}{ccc|c} 1 & -3 & 7 & 2 \\ 0 & 10 & -17 & -5 \\ 0 & 0 & 3 & 9 \end{array}\right). \tag{3.4.11}$$

Thus, elementary row operations transformed Equation 3.4.5 through Equation 3.4.7 into the triangular system:

$$x_1 - 3x_2 + 7x_3 = 2, \tag{3.4.12}$$

$$10x_2 - 17x_3 = -5, \tag{3.4.13}$$

$$3x_3 = 9, \tag{3.4.14}$$

which is *equivalent* to the original system. The final solution is obtained by *back substitution*, solving from Equation 3.4.14 back to Equation 3.4.12. In the present case, $x_3 = 3$. Then, $10x_2 = 17(3) - 5$, or $x_2 = 4.6$. Finally, $x_1 = 3x_2 - 7x_3 + 2 = -5.2$.

In general, if an $n \times n$ linear system can be reduced to triangular form, then it has a unique solution that we can obtain by performing back substitution. This reduction involves $n-1$ steps. In the first step, a pivot element, and thus the pivotal row, is chosen from the nonzero entries in the first column of the matrix. We interchange rows (if necessary) so that the pivotal row is the first row. Multiples of the pivotal row are then subtracted from each of the remaining $n-1$ rows so that there are 0's in the $(2,1), ..., (n,1)$ positions. In the second step, a pivot element is chosen from the nonzero entries in column 2, rows 2 through n, of the matrix. The row containing the pivot is then interchanged with the second row (if necessary) of the matrix and is used as the pivotal row. Multiples of the pivotal row are then subtracted from the remaining $n-2$ rows, eliminating all entries below the diagonal in the second column. The same procedure is repeated for columns 3 through $n-1$. Note that in the second step, row 1 and column 1 remain unchanged, in the third step the first two rows and first two columns remain unchanged, and so on.

If elimination is carried out as described, we arrive at an equivalent upper triangular system after $n-1$ steps. However, the procedure fails if, at any step, all possible choices for a pivot element equal zero. Let us now examine such cases.

Consider now the system

$$x_1 + 2x_2 + x_3 = -1, \tag{3.4.15}$$

$$2x_1 + 4x_2 + 2x_3 = -2, \tag{3.4.16}$$

$$x_1 + 4x_2 + 2x_3 = 2. \tag{3.4.17}$$

Its augmented matrix is

$$B = \left(\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 2 & 4 & 2 & -2 \\ 1 & 4 & 2 & 2 \end{array}\right). \tag{3.4.18}$$

Choosing the first row as our pivotal row, we find that

$$B = \left(\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 3 \end{array}\right), \tag{3.4.19}$$

or

$$B = \left(\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right). \tag{3.4.20}$$

The difficulty here is the presence of the zeros in the third row. Clearly any finite numbers satisfy the equation $0x_1 + 0x_2 + 0x_3 = 0$ and we have an infinite number of solutions. Closer examination of the original system shows an underdetermined system; Equation 3.4.15 and Equation 3.4.16 differ by a multiplicative factor of 2. An important aspect of this problem is the fact that the final augmented matrix is of the form of a staircase or *echelon form* rather than of triangular form.

Let us modify Equation 3.4.15 through Equation 3.4.17 to read

$$x_1 + 2x_2 + x_3 = -1, \tag{3.4.21}$$

$$2x_1 + 4x_2 + 2x_3 = 3, \tag{3.4.22}$$

$$x_1 + 4x_2 + 2x_3 = 2, \tag{3.4.23}$$

then the final augmented matrix is

$$B = \left(\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 5 \end{array}\right). \tag{3.4.24}$$

We again have a problem with the third row because $0x_1+0x_2+0x_3=5$, which is impossible. There is no solution in this case and we have an *inconsistent system.* Note, once again, that our augmented matrix has a row echelon form rather than a triangular form.

In summary, to include all possible situations in our procedure, we must rewrite the augmented matrix in row echelon form. We have *row echelon form* when:

(1) The first nonzero entry in each row is 1.

(2) If row k does not consist entirely of zeros, the number of leading zero entries in row $k+1$ is greater than the number of leading zero entries in row k.

(3) If there are rows whose entries are all zero, they are below the rows having nonzero entries.

The number of nonzero rows in the row echelon form of a matrix is known as its *rank.* In MATLAB, the rank is easily found using the command `rank( )`. *Gaussian elimination* is the process of using elementary row operations to transform a linear system into one whose augmented matrix is in row echelon form.

• Example 3.4.3

Each of the following matrices is *not* of row echelon form because they violate one of the conditions for row echelon form:

$$\begin{pmatrix} 2 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.4.25}$$

□

• Example 3.4.4

The following matrices are in row echelon form:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 4 & 6 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 4 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.4.26}$$

□

• Example 3.4.5

Gaussian elimination can also be used to solve the general problem $AX = B$. One of the most common applications is in finding the inverse. For example, let us find the inverse of the matrix

$$A = \begin{pmatrix} 4 & -2 & 2 \\ -2 & -4 & 4 \\ -4 & 2 & 8 \end{pmatrix} \tag{3.4.27}$$

by Gaussian elimination.

Because the inverse is defined by $AA^{-1} = I$, our augmented matrix is

$$\left(\begin{array}{ccc|ccc} 4 & -2 & 2 & 1 & 0 & 0 \\ -2 & -4 & 4 & 0 & 1 & 0 \\ -4 & 2 & 8 & 0 & 0 & 1 \end{array}\right). \quad (3.4.28)$$

Then, by elementary row operations,

$$\left(\begin{array}{ccc|ccc} 4 & -2 & 2 & 1 & 0 & 0 \\ -2 & -4 & 4 & 0 & 1 & 0 \\ -4 & 2 & 8 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{ccc|ccc} -2 & -4 & 4 & 0 & 1 & 0 \\ 4 & -2 & 2 & 1 & 0 & 0 \\ -4 & 2 & 8 & 0 & 0 & 1 \end{array}\right) \quad (3.4.29)$$

$$= \left(\begin{array}{ccc|ccc} -2 & -4 & 4 & 0 & 1 & 0 \\ 4 & -2 & 2 & 1 & 0 & 0 \\ 0 & 0 & 10 & 1 & 0 & 1 \end{array}\right) \quad (3.4.30)$$

$$= \left(\begin{array}{ccc|ccc} -2 & -4 & 4 & 0 & 1 & 0 \\ 0 & -10 & 10 & 1 & 2 & 0 \\ 0 & 0 & 10 & 1 & 0 & 1 \end{array}\right) \quad (3.4.31)$$

$$= \left(\begin{array}{ccc|ccc} -2 & -4 & 4 & 0 & 1 & 0 \\ 0 & -10 & 0 & 0 & 2 & -1 \\ 0 & 0 & 10 & 1 & 0 & 1 \end{array}\right) \quad (3.4.32)$$

$$= \left(\begin{array}{ccc|ccc} -2 & -4 & 0 & -2/5 & 1 & -2/5 \\ 0 & -10 & 0 & 0 & 2 & -1 \\ 0 & 0 & 10 & 1 & 0 & 1 \end{array}\right) \quad (3.4.33)$$

$$= \left(\begin{array}{ccc|ccc} -2 & 0 & 0 & -2/5 & 1/5 & 0 \\ 0 & -10 & 0 & 0 & 2 & -1 \\ 0 & 0 & 10 & 1 & 0 & 1 \end{array}\right) \quad (3.4.34)$$

$$= \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/5 & -1/10 & 0 \\ 0 & 1 & 0 & 0 & -1/5 & 1/10 \\ 0 & 0 & 1 & 1/10 & 0 & 1/10 \end{array}\right). \quad (3.4.35)$$

Thus, the right half of the augmented matrix yields the inverse and it equals

$$A^{-1} = \begin{pmatrix} 1/5 & -1/10 & 0 \\ 0 & -1/5 & 1/10 \\ 1/10 & 0 & 1/10 \end{pmatrix}. \quad (3.4.36)$$

MATLAB has the ability of doing Gaussian elimination step by step. We begin by typing

```
>>% input augmented matrix
>>aug = [4 -2 2 1 0 0 ; -2 -4 4 0 1 0;-4 2 8 0 0 1];
>>rrefmovie(aug);
```

The MATLAB command `rrefmovie(A)` produces the reduced row echelon form of A. Repeated pressing of any key gives the next step in the calculation along with a statement of how it computed the modified augmented matrix. Eventually you obtain

```
A =
            1           0           0         1/5       -1/10           0
```

```
        0         1         0         0      -1/5      1/10
        0         0         1      1/10         0      1/10
```

You can read the inverse matrix just as we did earlier. □

Gaussian elimination may be used with overdetermined systems. *Overdetermined systems* are linear systems where there are more equations than unknowns ($m > n$). These systems are usually (but not always) inconsistent.

• Example 3.4.6

Consider the linear system

$$x_1 + x_2 = 1, \tag{3.4.37}$$

$$-x_1 + 2x_2 = -2, \tag{3.4.38}$$

$$x_1 - x_2 = 4. \tag{3.4.39}$$

After several row operations, the augmented matrix

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ -1 & 2 & -2 \\ 1 & -1 & 4 \end{array}\right) \tag{3.4.40}$$

becomes

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -7 \end{array}\right). \tag{3.4.41}$$

From the last row of the augmented matrix, Equation 3.4.41, we see that the system is inconsistent.

If we test this system using MATLAB by typing

```
>>% input augmented matrix
>>aug = [1 1 1 ; -1 2 -2; 1 -1 4];
>>rrefmovie(aug);
```

eventually you obtain

```
A =
                     1                 0                 0
                     0                 1                 0
                     0                 0                 1
```

Although the numbers have changed from our hand calculation, we still have an inconsistent system because $x_1 = x_2 = 0$ does not satisfy $x_1 + x_2 = 1$.

Considering now a slight modification of this system to

$$x_1 + x_2 = 1, \tag{3.4.42}$$

$$-x_1 + 2x_2 = 5, \tag{3.4.43}$$

$$x_1 = -1, \tag{3.4.44}$$

the final form of the augmented matrix is

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right), \tag{3.4.45}$$

which has the unique solution $x_1 = -1$ and $x_2 = 2$.

How does MATLAB handle this problem? Typing

```
>>% input augmented matrix
>>aug = [1 1 1 ; -1 2 5; 1 0 -1];
>>rrefmovie(aug);
```

we eventually obtain

```
A =

     1     0    -1
     0     1     2
     0     0     0
```

This yields $x_1 = -1$ and $x_2 = 2$, as we found by hand.

Finally, by introducing the set:

$$x_1 + x_2 = 1, \tag{3.4.46}$$
$$2x_1 + 2x_2 = 2, \tag{3.4.47}$$
$$3x_1 + 3x_3 = 3, \tag{3.4.48}$$

the final form of the augmented matrix is

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right). \tag{3.4.49}$$

There are an infinite number of solutions: $x_1 = 1 - \alpha$, and $x_2 = \alpha$.

Turning to MATLAB, we first type

```
>>% input augmented matrix
>>aug = [1 1 1 ; 2 2 2; 3 3 3];
>>rrefmovie(aug);
```

and we eventually obtain

```
A =

     1     1     1
     0     0     0
     0     0     0
```

This is the same as Equation 3.4.49 and the final answer is the same. □

Gaussian elimination can also be employed with underdetermined systems. An *underdetermined linear system* is one where there are fewer equations than unknowns ($m < n$). These systems usually have an infinite number of solutions although they can be inconsistent.

• Example 3.4.7

Consider the underdetermined system:

$$2x_1 + 2x_2 + x_3 = -1, \tag{3.4.50}$$
$$4x_1 + 4x_2 + 2x_3 = 3. \tag{3.4.51}$$

Its augmented matrix can be transformed into the form:

$$\left(\begin{array}{ccc|c} 2 & 2 & 1 & -1 \\ 0 & 0 & 0 & 5 \end{array}\right). \tag{3.4.52}$$

Clearly this case corresponds to an inconsistent set of equations. On the other hand, if Equation 3.4.51 is changed to

$$4x_1 + 4x_2 + 2x_3 = -2, \tag{3.4.53}$$

then the final form of the augmented matrix is

$$\left(\begin{array}{ccc|c} 2 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{3.4.54}$$

and we have an infinite number of solutions, namely $x_3 = \alpha$, $x_2 = \beta$, and $2x_1 = -1-\alpha-2\beta$.□

Consider now one of the most important classes of linear equations: the homogeneous equations $A\mathbf{x} = \mathbf{0}$. If $\det(A) \neq 0$, then by Cramer's rule $x_1 = x_2 = x_3 = \cdots = x_n = 0$. Thus, the only possibility for a nontrivial solution is $\det(A) = 0$. In this case, A is singular, no inverse exists, and nontrivial solutions exist but they are not unique.

• Example 3.4.8

Consider the two homogeneous equations:

$$x_1 + x_2 = 0, \tag{3.4.55}$$
$$x_1 - x_2 = 0. \tag{3.4.56}$$

Note that $\det(A) = -2$. Solving this system yields $x_1 = x_2 = 0$.

However, if we change the system to

$$x_1 + x_2 = 0, \tag{3.4.57}$$
$$x_1 + x_2 = 0, \tag{3.4.58}$$

which has the $\det(A) = 0$ so that A is singular. Both equations yield $x_1 = -x_2 = \alpha$, any constant. Thus, there is an infinite number of solutions for this set of homogeneous equations. □

We close this section by outlining the proof of the theorem, which we introduced at the beginning.

Consider the system $A\mathbf{x} = \mathbf{b}$. By elementary row operations, the first equation in this system can be reduced to

$$x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n = \beta_1. \tag{3.4.59}$$

The second equation has the form

$$x_p + \alpha_{2p+1}x_{p+1} + \cdots + \alpha_{2n}x_n = \beta_2, \tag{3.4.60}$$

where $p > 1$. The third equation has the form

$$x_q + \alpha_{3q+1}x_{q+1} + \cdots + \alpha_{3n}x_n = \beta_3, \tag{3.4.61}$$

where $q > p$, and so on. To simplify the notation, we introduce z_i where we choose the first k values so that $z_1 = x_1$, $z_2 = x_p$, $z_3 = x_q$, ... Thus, the question of the existence of solutions depends upon the three integers: m, n, and k. The resulting set of equations have the form:

$$\begin{pmatrix} 1 & \gamma_{12} & \cdots & \gamma_{1k} & \gamma_{1k+1} & \cdots & \gamma_{1n} \\ 0 & 1 & \cdots & \gamma_{2k} & \gamma_{2k+1} & \cdots & \gamma_{2n} \\ & & & \vdots & & & \\ 0 & 0 & \cdots & 1 & \gamma_{kk+1} & \cdots & \gamma_{kn} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & & & \vdots & & & \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \\ z_{k+1} \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ \beta_{k+1} \\ \vdots \\ \beta_m \end{pmatrix}. \tag{3.4.62}$$

Note that β_{k+1}, ..., β_m need not be all zero.

There are three possibilities:

(a) $k < m$ and at least one of the elements $\beta_{k+1}, \ldots, \beta_m$ is nonzero. Suppose that an element β_p is nonzero ($p > k$). Then the pth equation is

$$0z_1 + 0z_2 + \cdots + 0z_n = \beta_p \neq 0. \tag{3.4.63}$$

However, this is a contradiction and the equations are inconsistent.

(b) $k = n$ and either (i) $k < m$ and all of the elements $\beta_{k+1}, \ldots, \beta_m$ are zero, or (ii) $k = m$. Then the equations have a unique solution that can be obtained by back-substitution.

(c) $k < n$ and either (i) $k < m$ and all of the elements $\beta_{k+1}, \ldots, \beta_m$ are zero, or (ii) $k = m$. Then, arbitrary values can be assigned to the $n - k$ variables $z_{k+1}, \ldots, z_n$. The equations can be solved for $z_1, z_2, \ldots, z_k$ and there is an infinity of solutions.

For homogeneous equations $\mathbf{b} = \mathbf{0}$, all of the β_i are zero. In this case, we have only two cases:

(b′) $k = n$, then Equation 3.4.62 has the solution $\mathbf{z} = \mathbf{0}$, which leads to the trivial solution for the original system $A\mathbf{x} = \mathbf{0}$.

(c′) $k < n$, the equations possess an infinity of solutions given by assigning arbitrary values to $z_{k+1}, \ldots, z_n$.

Problems

Solve the following systems of linear equations by Gaussian elimination. Check your answer using MATLAB.

1. $2x_1 + x_2 = 4,$ $\quad 5x_1 - 2x_2 = 1$
2. $x_1 + x_2 = 0,$ $\quad 3x_1 - 4x_2 = 1$
3. $-x_1 + x_2 + 2x_3 = 0,$ $\quad 3x_1 + 4x_2 + x_3 = 0,$ $\quad -x_1 + x_2 + 2x_3 = 0$

4. $4x_1 + 6x_2 + x_3 = 2,\quad 2x_1 + x_2 - 4x_3 = 3,\quad 3x_1 - 2x_2 + 5x_3 = 8$
5. $3x_1 + x_2 - 2x_3 = -3,\quad x_1 - x_2 + 2x_3 = -1,\quad -4x_1 + 3x_2 - 6x_3 = 4$
6. $x_1 - 3x_2 + 7x_3 = 2,\quad 2x_1 + 4x_2 - 3x_3 = -1,\quad -3x_1 + 7x_2 + 2x_3 = 3$
7. $x_1 - x_2 + 3x_3 = 5,\quad 2x_1 - 4x_2 + 7x_3 = 7,\quad 4x_1 - 9x_2 + 2x_3 = -15$
8. $x_1 + x_2 + x_3 + x_4 = -1,\quad 2x_1 - x_2 + 3x_3 = 1,\quad 2x_2 + 3x_4 = 15,\quad -x_1 + 2x_2 + x_4 = -2$

Find the inverse of each of the following matrices by Gaussian elimination. Check your answers using MATLAB.

9. $\begin{pmatrix} -3 & 5 \\ 2 & 1 \end{pmatrix}$ 10. $\begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$ 11. $\begin{pmatrix} 19 & 2 & -9 \\ -4 & -1 & 2 \\ -2 & 0 & 1 \end{pmatrix}$ 12. $\begin{pmatrix} 1 & 2 & 5 \\ 0 & -1 & 2 \\ 2 & 4 & 11 \end{pmatrix}$

13. Does $(A^2)^{-1} = (A^{-1})^2$? Justify your answer.

Project: Construction of a Finite Fourier Series

In Example 5.1.1 we show that the function $f(t)$ given by Equation 5.1.8 can be reexpressed

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + b_n \sin(nt), \qquad -\pi < t < \pi,$$

if

$$a_0 = \frac{\pi}{2}, \quad a_n = \frac{(-1)^n - 1}{n^2\pi}, \quad \text{and} \quad b_n = \frac{(-1)^{n+1}}{n}.$$

There we stated the Fourier series fits $f(t)$ in a "least squares sense." In Section 5.7 we will show that we could approximate $f(t)$ with the finite Fourier series

$$f(t) = \tfrac{1}{2}A_0 + \sum_{k=1}^{M-1} A_k \cos(kt) + B_k \sin(kt) + \tfrac{1}{2}A_M \cos(Mt),$$

if we sample $f(t)$ at $t_m = (2m + 1 - M)\pi/M$, where $m = 0, 1, 2, \ldots, M-1$ and M is an even integer. Then we can use Equation 5.7.12 and Equation 5.7.13 to compute A_k and B_k. Because MATLAB solves linear equations in a least-squares sense, this suggests that we could use MATLAB as an alternative method for finding a finite Fourier series approximation.

Let us assume that

$$f(t) = \frac{A_0}{2} + \sum_{n=1}^{N} A_n \cos(nt) + B_n \sin(nt).$$

Then sampling $f(t)$ at the temporal points $t_m = -\pi + (2m-1)\pi/M$, we obtain the following system of linear equations:

$$\frac{A_0}{2} + \sum_{n=1}^{N} A_n \cos(nt_m) + B_n \sin(nt_m) = f(t_m),$$

where $m = 1, 2, \ldots, M$. Write a MATLAB program that solves this system for given N and M and compare your results with the exact answers a_0, a_n and b_n for various N and M.

Consider the case when $M > 2N+1$ (overspecified system), $M < 2N+1$ (underspecified system), and $M = 2N+1$ (equal number of unknowns and equations). Does this method yield any good results? If so, under which conditions?

Project: Solving Fredholm Integral Equation of the Second Kind

Fredholm equations of the second kind and their variants appear in many scientific and engineering applications. In this project you will use matrix methods to solve this equation:

$$u(x) = \int_a^b K(x,t)u(t)\,dt + f(x), \qquad a < x < b,$$

where the kernel $K(x,t)$ is a given real-valued and continuous function and $u(x)$ is the unknown. One method for solving this integral equation replaces the integral with some grid-point representation. The goal of this project is examine how we can use linear algebra to solve this numerical approximation to Fredholm's integral equation.

Step 1: Using Simpson's rule, show that our Fredholm equation can be written in the matrix form $(I - KD)\mathbf{u} = \mathbf{f}$, where

$$D = \begin{pmatrix} A_0 & 0 & \cdots & 0 & 0 \\ 0 & A_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & A_n \end{pmatrix},$$

$$K = \begin{pmatrix} K(x_0,x_0) & K(x_0,x_1) & \cdots & K(x_0,x_{n-1}) & K(x_0,x_n) \\ K(x_1,x_0) & K(x_1,x_1) & \cdots & K(x_1,x_{n-1}) & K(x_1,x_n) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ K(x_{n-1},x_0) & K(x_{n-1},x_1) & \cdots & K(x_{n-1},x_{n-1}) & K(x_{n-1},x_n) \\ K(x_n,x_0) & K(x_n,x_1) & \cdots & K(x_n,x_{n-1}) & K(x_n,x_n) \end{pmatrix},$$

$$\mathbf{u} = [u(x_0), u(x_1), \cdots, u(x_n)]^T, \quad \text{and} \quad \mathbf{f} = [f(x_0), f(x_1), \cdots, f(x_n)]^T,$$

where $A_0 = A_n = h/3$, $A_2 = A_4 = \cdots = A_{n-2} = 2h/3$, $A_1 = A_3 = \cdots = A_{n-1} = 4h/3$, $x_i = ih$, and $h = (b-a)/n$. Here n must be an *even* integer.

Step 2: Use MATLAB to solve our matrix equation to find $\mathbf{u}$. Use the following known solutions:

(a)	$K(x,t) = \frac{1}{2}x^2t^2,$	$f(x) = 0.9x^2,$	$u(x) = x^2,$
(b)	$K(x,t) = x^2e^{t(x-1)},$	$f(x) = x + (1-x)e^x,$	$u(x) = e^x,$
(c)	$K(x,t) = \frac{1}{3}e^{x-t},$	$f(x) = \frac{2}{3}e^x,$	$u(x) = e^x,$
(d)	$K(x,t) = -\frac{1}{3}e^{2x-5t/3},$	$f(x) = e^2x + \frac{1}{3},$	$u(x) = e^{2x},$
(e)	$K(x,t) = -x\left(e^{xt} - 1\right),$	$f(x) = e^x - x,$	$u(x) = 1,$
(f)	$K(x,t) = \frac{1}{2}xt,$	$f(x) = \frac{5}{6}x,$	$u(x) = x,$

when $0 \le x \le 1$. How does the accuracy of this method vary with n (or h)? What happens when n becomes large? Figure 3.4.1 shows the absolute value of the relative error in the

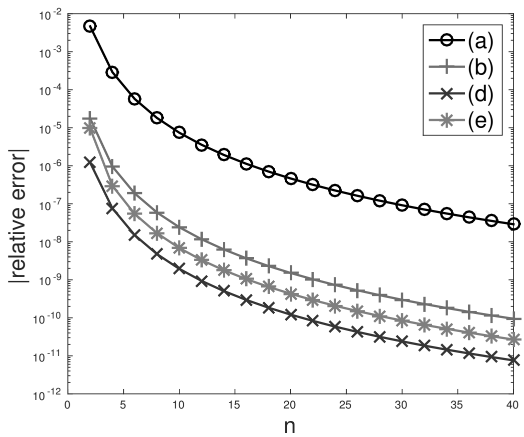


Figure 3.4.1: The absolute value of the relative error of the numerical solution of a Fredholm integral equation of the second kind as a function of n for test problems (a), (b), (d), and (e).

numerical solution at $x = \frac{1}{2}$ as a function of n. For test cases (c) and (f) the error was the same order of magnitude as the round-off error.

Project: LU Decomposition

In this section we showed how Gaussian elimination can be used to find solutions to sets of linear equations. A popular alternative involves rewriting the $n \times n$ coefficient matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_{ij} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \cdot & \cdot & \cdot & a_{nn} \end{pmatrix}$$

as the product of a lower $n \times n$ triangular matrix:

$$L = \begin{pmatrix} \ell_{11} & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \ell_{21} & \ell_{22} & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \ell_{ij} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdot & \cdot & \cdot & \ell_{nn} \end{pmatrix}$$

and an upper $n \times n$ triangular matrix:

$$U = \begin{pmatrix} 1 & u_{12} & u_{13} & \cdot & \cdot & \cdot & u_{1n} \\ 0 & 1 & u_{23} & \cdot & \cdot & \cdot & u_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 1 \end{pmatrix},$$

so that $A = LU$. By simply doing the matrix multiplication, we find the following Crout algorithm to compute ℓ_{ij} and u_{ij}:

$$\ell_{ij} = a_{ij} - \sum_{k=1}^{j-1} \ell_{ik} u_{kj}, \qquad j \leq i, \quad i = 1, 2, \ldots, n;$$

and

$$u_{ij} = \left[a_{ij} - \sum_{k=1}^{i-1} \ell_{ik} u_{kj} \right] \bigg/ \ell_{ii}, \qquad i < j, \quad j = 2, 3, \ldots, n.$$

For the special case of $j = 1$, $\ell_{i1} = a_{i1}$; for $i = 1$, $u_{1j} = a_{1j}/\ell_{11} = a_{i1}/a_{11}$. Clearly we could write code to compute L and U given A. However, MATLAB has a subroutine for doing this factorization `[L,U] = lu(A)`.

How does this factorization help us to solve $A\mathbf{x} = \mathbf{b}$? The goal of this project is to answer this question.

Step 1: Show that $L\mathbf{y} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{y}$ can be combined together to yield $A\mathbf{x} = \mathbf{b}$.

Step 2: Show that y_i and x_i can be computed from $y_1 = b_1/\ell_{11}$,

$$y_i = \left[b_i - \sum_{j=1}^{i-1} \ell_{ij} b_j \right] \bigg/ \ell_{ii}, \qquad i = 2, 3, \ldots, n;$$

and $x_n = y_n/u_{nn}$,

$$x_i = \left[y_i - \sum_{j=i+1}^{n} u_{ij} y_j \right] \bigg/ u_{ii}, \qquad i = n-1, n-2, \ldots, 1.$$

Step 3: Write a MATLAB script to solve $A\mathbf{x} = \mathbf{b}$ using LU decomposition.

Step 4: Check your program by resolving Problems 4, 6, 7, and 8.

The principal reason that this scheme is so popular is its ecomony of storage. The 0's in either L or U are not stored. Furthermore, after the element a_{ij} is used, it never appears again.

Project: QR Decomposition

In the previous project you discovered that by factoring the matrix A into upper and lower diagonal matrices, we could solve $A\mathbf{x} = \mathbf{b}$. Here we will again factor the matrix A into the product QR but Q will have the property that $Q^TQ = I$ (orthogonal matrix) and R is an upper triangular matrix.

Step 1: Assuming that we can rewrite $A\mathbf{x} = \mathbf{b}$ as $QR\mathbf{x} = \mathbf{b}$, multiply both sides of this second equation by Q^T and show that you obtain $R\mathbf{x} = Q^T\mathbf{b} = \mathbf{y}$.

Step 2: Show that x_i can be computed from $x_n = y_n/r_{nn}$ and

$$x_i = \left[y_i - \sum_{j=i+1}^{n} r_{ij} y_j \right] \Bigg/ r_{ii}, \qquad i = n-1, n-2, \ldots, 1.$$

Step 3: Write a MATLAB script to solve $A\mathbf{x} = \mathbf{b}$ using QR decomposition.

Step 4: Check your program by resolving Problems 4, 6, 7, and 8.

What advantages does QR decomposition have over LU decomposition? First, solving $A\mathbf{x} = \mathbf{b}$ via $R\mathbf{x} = Q^T\mathbf{b}$ is as well-conditioned as the original problem. Second, QR decomposition finds the least-squares solutions when no exact solution exists. When there are exact solutions, it finds all of them.

3.5 EIGENVALUES AND EIGENVECTORS

One of the classic problems of linear algebra[6] is finding all of the λ's that satisfy the $n \times n$ system

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{3.5.1}$$

The nonzero quantity λ is the *eigenvalue* or *characteristic value* of A. The vector $\mathbf{x}$ is the *eigenvector* or *characteristic vector* belonging to λ. The set of the eigenvalues of A is called the *spectrum* of A. The largest of the absolute values of the eigenvalues of A is called the *spectral radius* of A.

To find λ and $\mathbf{x}$, we first rewrite Equation 3.5.1 as a set of homogeneous equations:

$$(A - \lambda I)\mathbf{x} = \mathbf{0}. \tag{3.5.2}$$

From the theory of linear equations, Equation 3.5.2 has trivial solutions unless its determinant equals zero. On the other hand, if

$$\det(A - \lambda I) = 0, \tag{3.5.3}$$

there are an infinity of solutions.

The expansion of the determinant, Equation 3.5.3, yields an nth-degree polynomial in λ, the *characteristic polynomial*. The roots of the characteristic polynomial are the eigenvalues of A. Because the characteristic polynomial has exactly n roots, A has n eigenvalues, some of which can be repeated (with multiplicity $k \le n$) and some of which can be complex numbers. For each eigenvalue λ_i, there is a corresponding eigenvector $\mathbf{x}_i$. This eigenvector is the solution of the homogeneous equations $(A - \lambda_i I)\mathbf{x}_i = \mathbf{0}$.

An important property of eigenvectors is their *linear independence* if there are n distinct eigenvalues. Vectors are linearly independent if the equation

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n = \mathbf{0} \tag{3.5.4}$$

can be satisfied only by taking *all* of the coefficients α_n equal to zero.

---

[6] The standard reference is Wilkinson, J. H., 1965: *The Algebraic Eigenvalue Problem*. Oxford University Press, 662 pp.

This concept of linear independence or dependence actually extends to vectors in general, not just eigenvectors. Algebraists would say that our n linearly independent vectors form a *basis* that *spans* a *vector space* V. A vector space is simply a set V of vectors that can be added and scaled. The maximum number of linearly independent vectors in a vector space gives its *dimension* of V. A vector space V can have many different bases, but there are always the same number of basis vectors in each of them.

Returning to the eigenvalue problem, we now show that in the case of n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, each eigenvalue λ_i having a corresponding eigenvector $\mathbf{x}_i$, the eigenvectors form a basis. We first write down the linear dependence condition

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_n\mathbf{x}_n = \mathbf{0}. \tag{3.5.5}$$

Premultiplying Equation 3.5.5 by A,

$$\alpha_1 A\mathbf{x}_1 + \alpha_2 A\mathbf{x}_2 + \cdots + \alpha_n A\mathbf{x}_n = \alpha_1\lambda_1\mathbf{x}_1 + \alpha_2\lambda_2\mathbf{x}_2 + \cdots + \alpha_n\lambda_n\mathbf{x}_n = \mathbf{0}. \tag{3.5.6}$$

Premultiplying Equation 3.5.5 by A^2,

$$\alpha_1 A^2\mathbf{x}_1 + \alpha_2 A^2\mathbf{x}_2 + \cdots + \alpha_n A^2\mathbf{x}_n = \alpha_1\lambda_1^2\mathbf{x}_1 + \alpha_2\lambda_2^2\mathbf{x}_2 + \cdots + \alpha_n\lambda_n^2\mathbf{x}_n = \mathbf{0}. \tag{3.5.7}$$

In a similar manner, we obtain the system of equations:

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} \alpha_1\mathbf{x}_1 \\ \alpha_2\mathbf{x}_2 \\ \alpha_3\mathbf{x}_3 \\ \vdots \\ \alpha_n\mathbf{x}_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{3.5.8}$$

Because

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{vmatrix} = \begin{matrix} (\lambda_2-\lambda_1)(\lambda_3-\lambda_2)(\lambda_3-\lambda_1)(\lambda_4-\lambda_3) \\ (\lambda_4-\lambda_2)\cdots(\lambda_n-\lambda_1) \neq 0, \end{matrix} \tag{3.5.9}$$

since it is a Vandermonde determinant, $\alpha_1\mathbf{x}_1 = \alpha_2\mathbf{x}_2 = \alpha_3\mathbf{x}_3 = \cdots = \alpha_n\mathbf{x}_n = 0$. Because the eigenvectors are nonzero, $\alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$, and the eigenvectors are linearly independent. □

This property of eigenvectors allows us to express any arbitrary vector $\mathbf{x}$ as a linear sum of the eigenvectors $\mathbf{x}_i$, or

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n. \tag{3.5.10}$$

We will make good use of this property in Example 3.5.3.

• **Example 3.5.1**

Let us find the eigenvalues and corresponding eigenvectors of the matrix

$$A = \begin{pmatrix} -4 & 2 \\ -1 & -1 \end{pmatrix}. \tag{3.5.11}$$

We begin by setting up the characteristic equation:

$$\det(A - \lambda I) = \begin{vmatrix} -4-\lambda & 2 \\ -1 & -1-\lambda \end{vmatrix} = 0. \qquad (3.5.12)$$

Expanding the determinant,

$$(-4-\lambda)(-1-\lambda) + 2 = \lambda^2 + 5\lambda + 6 = (\lambda+3)(\lambda+2) = 0. \qquad (3.5.13)$$

Thus, the eigenvalues of the matrix A are $\lambda_1 = -3$, and $\lambda_2 = -2$.

To find the corresponding eigenvectors, we must solve the linear system:

$$\begin{pmatrix} -4-\lambda & 2 \\ -1 & -1-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \qquad (3.5.14)$$

For example, for $\lambda_1 = -3$,

$$\begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad (3.5.15)$$

or

$$x_1 = 2x_2. \qquad (3.5.16)$$

Thus, any nonzero multiple of the vector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ is an eigenvector belonging to $\lambda_1 = -3$. Similarly, for $\lambda_2 = -2$, the eigenvector is any nonzero multiple of the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Of course, MATLAB will do all of the computations for you via the command `eig`, which computes the eigenvalues and corresponding eigenvalues. In the present case, you would type

```
>> A = [-4 2; -1 -1]; % load in array A
>> % find eigenvalues and eigenvectors
>> [eigenvector,eigenvalue] = eig(A)
```

This yields

```
eigenvector =
                              -0.8944                  -0.7071
                              -0.4472                  -0.7071
```

and

```
eigenvalue =
                                 -3                       0
                                  0                      -2.
```

The eigenvalues are given as the elements along the principal diagonal of `eigenvalue`. The corresponding vectors are given by the corresponding column of `eigenvector`. As this example shows, these eigenvectors have been normalized so that their norm, Equation 3.1.5, equals one. Also, their sign may be different than any you would choose. We can recover our hand-computed results by dividing the first eigenvector by -0.4472 while in the second case we would divide by -0.7071. Finally, note that the product `eigenvector*eigenvalue*inv(eigenvector)` would yield `A`. □

• **Example 3.5.2**

Let us now find the eigenvalues and corresponding eigenvectors of the matrix

$$A = \begin{pmatrix} -4 & 5 & 5 \\ -5 & 6 & 5 \\ -5 & 5 & 6 \end{pmatrix}. \tag{3.5.17}$$

Setting up the characteristic equation:

$$\begin{aligned} \det(A - \lambda I) & \\ &= \begin{vmatrix} -4-\lambda & 5 & 5 \\ -5 & 6-\lambda & 5 \\ -5 & 5 & 6-\lambda \end{vmatrix} = \begin{vmatrix} -4-\lambda & 5 & 5 \\ -5 & 6-\lambda & 5 \\ 0 & \lambda-1 & 1-\lambda \end{vmatrix} \quad (3.5.18) \\ &= (\lambda-1)\begin{vmatrix} -4-\lambda & 5 & 5 \\ -5 & 6-\lambda & 5 \\ 0 & 1 & -1 \end{vmatrix} = (\lambda-1)^2\begin{vmatrix} -1 & 1 & 0 \\ -5 & 6-\lambda & 5 \\ 0 & 1 & -1 \end{vmatrix} \quad (3.5.19) \end{aligned}$$

$$\det(A - \lambda I) = (\lambda-1)^2 \begin{vmatrix} -1 & 0 & 0 \\ -5 & 6-\lambda & 0 \\ 0 & 1 & -1 \end{vmatrix} = (\lambda-1)^2(6-\lambda) = 0. \tag{3.5.20}$$

Thus, the eigenvalues of the matrix A are $\lambda_{1,2} = 1$ (twice), and $\lambda_3 = 6$.

To find the corresponding eigenvectors, we must solve the linear system:

$$(-4-\lambda)x_1 + 5x_2 + 5x_3 = 0, \tag{3.5.21}$$

$$-5x_1 + (6-\lambda)x_2 + 5x_3 = 0, \tag{3.5.22}$$

and

$$-5x_1 + 5x_2 + (6-\lambda)x_3 = 0. \tag{3.5.23}$$

For $\lambda_3 = 6$, Equations 3.5.21 through 3.5.23 become

$$-10x_1 + 5x_2 + 5x_3 = 0, \tag{3.5.24}$$

$$-5x_1 + 5x_3 = 0, \tag{3.5.25}$$

and

$$-5x_1 + 5x_2 = 0. \tag{3.5.26}$$

Thus, $x_1 = x_2 = x_3$ and the eigenvector is any nonzero multiple of the vector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

The interesting aspect of this example centers on finding the eigenvector for the eigenvalue $\lambda_{1,2} = 1$. If $\lambda_{1,2} = 1$, then Equations 3.5.21 through 3.5.23 collapse into one equation,

$$-x_1 + x_2 + x_3 = 0, \tag{3.5.27}$$

and we have *two* free parameters at our disposal. Let us take $x_2 = \alpha$, and $x_3 = \beta$. Then the eigenvector equals $\alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ for $\lambda_{1,2} = 1$.

In this example, we may associate the eigenvector $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ with $\lambda_1 = 1$, and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ with $\lambda_2 = 1$ so that, along with the eigenvector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ with $\lambda_3 = 6$, we still have n *linearly independent* eigenvectors for our 3×3 matrix. However, with repeated eigenvalues this is not always true. For example,

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \tag{3.5.28}$$

has the repeated eigenvalues $\lambda_{1,2} = 1$. However, there is only a single eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ for *both* λ_1 and λ_2.

What happens in MATLAB in the present case? Typing in

```
>> A = [-4 5 5; -5 6 5; -5 5 6]; % load in array A
>> % find eigenvalues and eigenvectors
>> [eigenvector,eigenvalue] = eig(A)
```

we obtain

```
eigenvector =
                   -0.8165              0.5774              0.4259
                   -0.4082              0.5774             -0.3904
                   -0.4082              0.5774              0.8162
```

and

```
eigenvalue =
                          1                 0                 0
                          0                 6                 0
                          0                 0                 1
```

The second eigenvector is clearly the same as the hand-computed one if you normalize it with 0.5774. The equivalence of the first and third eigenvectors is not as clear. However, if you choose $\alpha = \beta = -0.4082$, then the first eigenvector agrees with the hand-computed value. Similarly, taking $\alpha = -0.3904$ and $\beta = 0.8162$ result in agreement with the third MATLAB eigenvector. Finally, note that the product `eigenvector*eigenvalue*inv(eigenvector)` would yield `A`. □

• Example 3.5.3

When we discussed the stability of numerical schemes for the wave equation in Section 7.6, we will examine the behavior of a prototypical Fourier harmonic to variations in the parameter $c\Delta t/\Delta x$. In this example we shall show another approach to determining the stability of a numerical scheme via matrices.

Consider the explicit scheme for the numerical integration of the wave equation, Equation 7.6.11. We can rewrite that single equation as the coupled difference equations:

$$u_m^{n+1} = 2(1 - r^2)u_m^n + r^2(u_{m+1}^n + u_{m-1}^n) - v_m^n, \tag{3.5.29}$$

and

$$v_m^{n+1} = u_m^n, \tag{3.5.30}$$

where $r = c\Delta t/\Delta x$. Let $u_{m+1}^n = e^{i\beta\Delta x}u_m^n$, and $u_{m-1}^n = e^{-i\beta\Delta x}u_m^n$, where β is real. Then Equation 3.5.29 and Equation 3.5.30 become

$$u_m^{n+1} = 2\left[1 - 2r^2\sin^2\left(\frac{\beta\Delta x}{2}\right)\right]u_m^n - v_m^n, \qquad (\mathbf{3.5.31})$$

and

$$v_m^{n+1} = u_m^n, \qquad (\mathbf{3.5.32})$$

or in the matrix form

$$\mathbf{u}_m^{n+1} = \begin{pmatrix} 2\left[1 - 2r^2\sin^2\left(\frac{\beta\Delta x}{2}\right)\right] & -1 \\ 1 & 0 \end{pmatrix}\mathbf{u}_m^n, \qquad (\mathbf{3.5.33})$$

where $\mathbf{u}_m^n = \begin{pmatrix} u_m^n \\ v_m^n \end{pmatrix}$. The eigenvalues λ of this *amplification matrix* are given by

$$\lambda^2 - 2\left[1 - 2r^2\sin^2\left(\frac{\beta\Delta x}{2}\right)\right]\lambda + 1 = 0, \qquad (\mathbf{3.5.34})$$

or

$$\lambda_{1,2} = 1 - 2r^2\sin^2\left(\frac{\beta\Delta x}{2}\right) \pm 2r\sin\left(\frac{\beta\Delta x}{2}\right)\sqrt{r^2\sin^2\left(\frac{\beta\Delta x}{2}\right) - 1}. \qquad (\mathbf{3.5.35})$$

Because each successive time step consists of multiplying the solution from the previous time step by the amplification matrix, the solution is stable only if $\mathbf{u}_m^n$ remains bounded. This occurs only if all of the eigenvalues have a magnitude less or equal to one, because

$$\mathbf{u}_m^n = \sum_k c_k A^n \mathbf{x}_k = \sum_k c_k \lambda_k^n \mathbf{x}_k, \qquad (\mathbf{3.5.36})$$

where A denotes the amplification matrix and $\mathbf{x}_k$ denotes the eigenvectors corresponding to the eigenvalues λ_k. Equation 3.5.36 follows from our ability to express any initial condition in terms of an eigenvector expansion

$$\mathbf{u}_m^0 = \sum_k c_k \mathbf{x}_k. \qquad (\mathbf{3.5.37})$$

In our particular example, two cases arise. If $r^2\sin^2(\beta\Delta x/2) \le 1$,

$$\lambda_{1,2} = 1 - 2r^2\sin^2\left(\frac{\beta\Delta x}{2}\right) \pm 2ri\sin\left(\frac{\beta\Delta x}{2}\right)\sqrt{1 - r^2\sin^2\left(\frac{\beta\Delta x}{2}\right)} \qquad (\mathbf{3.5.38})$$

and $|\lambda_{1,2}| = 1$. On the other hand, if $r^2\sin^2(\beta\Delta x/2) > 1$, $|\lambda_{1,2}| > 1$. Thus, we have stability only if $c\Delta t/\Delta x \le 1$.

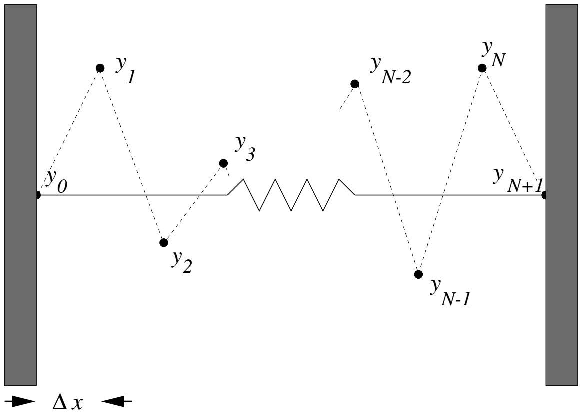


Figure 3.5.1: Schematic for finite-differencing a Sturm-Liouville problem into a set of difference equations.

Problems

Find the eigenvalues and corresponding eigenvectors for the following matrices. Check your answers using MATLAB.

1. $A = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}$
2. $A = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$
3. $A = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}$
4. $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$
5. $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$
6. $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 5 & -1 \end{pmatrix}$
7. $A = \begin{pmatrix} 4 & -5 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$
8. $A = \begin{pmatrix} -2 & 0 & 1 \\ 3 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$

Project: Numerical Solution of the Sturm-Liouville Problem

You may have been struck by the similarity of the algebraic eigenvalue problem to the Sturm-Liouville problem. (See Section 6.1.) In both cases nontrivial solutions exist only for characteristic values of λ. The purpose of this project is to further deepen your insight into these similarities.

Consider the Sturm-Liouville problem

$$y'' + \lambda y = 0, \qquad y(0) = y(\pi) = 0. \tag{3.5.39}$$

We know that it has the nontrivial solutions $\lambda_m = m^2$, $y_m(x) = \sin(mx)$, where $m = 1, 2, 3, \ldots$.

Step 1: Let us solve this problem numerically. Introducing centered finite differencing and the grid shown in Figure 3.5.1, show that

$$y'' \approx \frac{y_{n+1} - 2y_n + y_{n-1}}{(\Delta x)^2}, \qquad n = 1, 2, \ldots, N, \tag{3.5.40}$$

Table 3.5.1: Eigenvalues Computed from Equation 3.5.42 as a Numerical Approximation of the Sturm-Liouville Problem, Equation 3.5.39

N	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
1	0.81057						
2	0.91189	2.73567					
3	0.94964	3.24228	5.53491				
4	0.96753	3.50056	6.63156	9.16459			
5	0.97736	3.64756	7.29513	10.94269	13.61289		
6	0.98333	3.73855	7.71996	12.13899	16.12040	18.87563	
7	0.98721	3.79857	8.00605	12.96911	17.93217	22.13966	24.95100
8	0.98989	3.84016	8.20702	13.56377	19.26430	24.62105	28.98791
20	0.99813	3.97023	8.84993	15.52822	23.85591	33.64694	44.68265
50	0.99972	3.99498	8.97438	15.91922	24.80297	35.59203	48.24538

where $\Delta x = \pi/(N+1)$. Show that the finite-differenced form of Equation 3.5.39 is

$$-h^2 y_{n+1} + 2h^2 y_n - h^2 y_{n-1} = \lambda y_n \tag{3.5.41}$$

with $y_0 = y_{N+1} = 0$, and $h = 1/(\Delta x)$.

Step 2: Solve Equation 3.5.41 as an algebraic eigenvalue problem using $N = 1, 2, \ldots$. Show that Equation 3.5.41 can be written in the matrix form of

$$\begin{pmatrix} 2h^2 & -h^2 & 0 & \cdots & 0 & 0 & 0 \\ -h^2 & 2h^2 & -h^2 & \cdots & 0 & 0 & 0 \\ 0 & -h^2 & 2h^2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -h^2 & 2h^2 & -h^2 \\ 0 & 0 & 0 & \cdots & 0 & -h^2 & 2h^2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{N-1} \\ y_N \end{pmatrix} = \lambda \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{N-1} \\ y_N \end{pmatrix}. \tag{3.5.42}$$

Note that the coefficient matrix is symmetric.

Step 3: You are now ready to compute the eigenvalues. For small N this could be done by hand. However, it is easier just to write a MATLAB program that will handle any $N \geq 2$. Table 3.5.1 has been provided so that you can check your program.

With your program, answer the following questions: How do your computed eigenvalues compare to the eigenvalues given by the Sturm-Liouville problem? What happens as you increase N? Which computed eigenvalues agree best with those given by the Sturm-Liouville problem? Which ones compare the worst?

Step 4: Let us examine the eigenfunctions now. Starting with the smallest eigenvalue, use MATLAB to plot Cy_j as a function of x_i where y_j is the jth eigenvector, $j = 1, 2, \ldots, N$, $x_i = i\Delta x$, $i = 1, 2, \ldots, N$, and C is chosen so that $C^2 \Delta x \sum_i y_j^2(x_i) = 1$. On the same plot, graph $y_j(x) = \sqrt{2/\pi}\,\sin(jx)$. Why did we choose C as we did? Which eigenvectors and eigenfunctions agree the best? Which eigenvectors and eigenfunctions agree the worst? Why? Why are there N eigenvectors and an infinite number of eigenfunctions?

Step 5: The most important property of eigenfunctions is orthogonality. But what do we mean by orthogonality in the case of eigenvectors? Recall from three-dimensional vectors we had the scalar dot product

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3. \tag{3.5.43}$$

For n-dimensional vectors, this dot product is generalized to the inner product

$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k y_k. \tag{3.5.44}$$

Orthogonality implies that $\mathbf{x} \cdot \mathbf{y} = 0$ if $\mathbf{x} \neq \mathbf{y}$. Are your eigenvectors orthogonal? How might you use this property with eigenvectors?

Project: Singular Value Decomposition and Linear Least Squares

In the previous section we showed two ways that linear equations can be solved by factoring the matrix A in $A\mathbf{x} = \mathbf{b}$ into a product of two matrices. The LU and QR decompositions are *not* the only possible factorization. One popular version rewrites the square matrix A as PDP^{-1}, where D is a diagonal matrix with the n eigenvalues along the principal diagonal and P contains the eigenvectors in the transition matrix P. MATLAB's routine `eig` yields both D and P via `[P,D] = eig(A)`. See Example 3.6.4. In this project we focus on singular value decomposition, possibly the most important matrix decomposition of them all. It is used in signal processing, statistics, and numerical methods and theory.

Singular value decomposition factorizes a matrix A of dimension $m \times n$ into the product of three matrices: $A = UDV^T$, where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix, and D is an $m \times n$ diagonal matrix. The diagonal entries of D are called the *singular values* of A. The rank of a matrix equals the number of non-zero singular values.

Step 1: Consider the matrix A given by

$$A = \begin{pmatrix} 45 & -108 & 36 & -45 \\ 21 & -68 & 26 & -33 \\ 72 & -32 & -16 & 24 \\ -56 & 64 & -8 & 8 \\ 50 & -32 & -6 & 10 \end{pmatrix}.$$

Using MATLAB's subroutine `svd`, confirm that it can factorize the array A into U, D, and V. Then check that $A = UDV^T$ and find the rank of this matrix.

The goal of this project is to find the parameters m and c so that the line $y = mx + c$ gives the best fit to n data points. If there are only two data points, there is no problem because we could immediately find the slope m and the intercept c. However, if $n > 2$ we cannot hope to choose these coefficients so that the straight line fits them. How does singular value decomposition come to the rescue?

To find the answer, we begin by noting that each data point (x_i, y_i) must satisfy the linear equation $mx_i + c = y_i$, or

$$m \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + c \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} = A\mathbf{x} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \mathbf{y}.$$

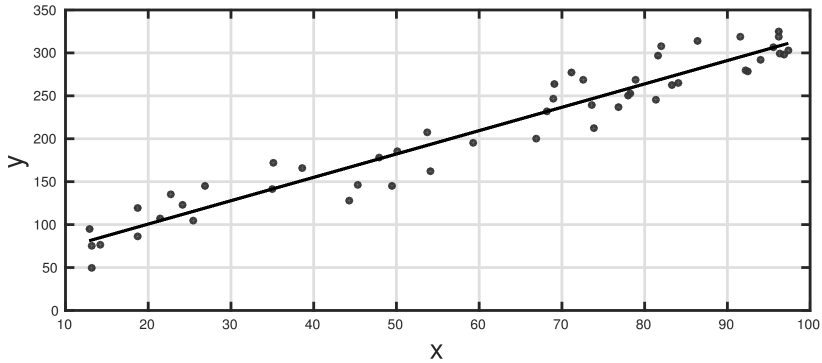


Figure 3.5.2: Using 50 "data points," the singular value decomposition provides a least-squares fit to the data.

We can write this over-determined system of linear equations $A\mathbf{x} = \mathbf{b}$. Let the residual $\mathbf{r}$ be a vector defined by $\mathbf{r} = A\mathbf{x} - \mathbf{b}$. The vector $\mathbf{x}^*$ that yields the smallest possible residual in the least-squares sense is $||\mathbf{r}|| = ||A\mathbf{x}^* - \mathbf{b}|| \leq ||A\mathbf{x} - \mathbf{b}||$, where $||\cdot||$ denotes the Euclidean norm. Although a least-squares solution always exists, it might not be unique. However, the least-squares solution $\mathbf{x}$ with the *smallest* norm $||\mathbf{x}||$ is unique and equals $A^TA\mathbf{x} = A^T\mathbf{b}$ or $\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$. This solution can be found using singular value decomposition as $\mathbf{x} = VD_0^{-1}U^T\mathbf{b}$, where the $n \times m$ matrix D_0^{-1} has the diagonal terms:

$$\left(D_0^{-1}\right)_{ii} = \begin{cases} 1/\Sigma_i, & \text{if } \Sigma_i > \epsilon, \\ 0, & \text{otherwise,} \end{cases}$$

even when A is singular or ill-conditioned.

Step 2: In Step 1, we found that the rank for the matrix A is 2. What values of $\mathbf{x}$ does singular value decomposition give if $\mathbf{b} = (\,-2\;4\;4\;6\;\;-4\,)^T$? This solution equals the least-squares solution of minimum length.

Step 3: Returning to our original goal of finding the best linear fit to data, create data for your numerical experiment. One way would use the simple line $y = mx + c$ (with arbitrary chosen values of m and c) to create the initial data and then use the random number generator `rand` to modify this initial data (both x and y) so that both have "noise."

Step 4: Construct the array A and column vector $\mathbf{y}$. Using the MATLAB routine `svd`, find $\mathbf{x} = VD_0^{-1}U^T\mathbf{b}$.

Step 5: Construct the least-squares fit for the data and plot this curve and your data on the same figure. See Figure 3.5.2.

3.6 SYSTEMS OF LINEAR DIFFERENTIAL EQUATIONS

In this section we show how we may apply the classic algebraic eigenvalue problem to solve a system of ordinary differential equations.

Let us solve the following system:

$$x_1' = x_1 + 3x_2, \tag{3.6.1}$$

and

$$x_2' = 3x_1 + x_2, \tag{3.6.2}$$

where the primes denote the time derivative.

We begin by rewriting Equation 3.6.1 and Equation 3.6.2 in matrix notation:

$$\mathbf{x}' = A\mathbf{x}, \tag{3.6.3}$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad \text{and} \qquad A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}. \tag{3.6.4}$$

Note that

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{x}'. \tag{3.6.5}$$

Assuming a solution of the form

$$\mathbf{x} = \mathbf{x}_0 e^{\lambda t}, \qquad \text{where} \qquad \mathbf{x}_0 = \begin{pmatrix} a \\ b \end{pmatrix} \tag{3.6.6}$$

is a constant vector, we substitute Equation 3.6.6 into Equation 3.6.3 and find that

$$\lambda e^{\lambda t}\mathbf{x}_0 = A e^{\lambda t}\mathbf{x}_0. \tag{3.6.7}$$

Because $e^{\lambda t}$ does not generally equal zero, we have that

$$(A - \lambda I)\mathbf{x}_0 = \mathbf{0}, \tag{3.6.8}$$

which we solved in the previous section. This set of homogeneous equations is the *classic eigenvalue problem.* In order for this set not to have trivial solutions,

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{vmatrix} = 0. \tag{3.6.9}$$

Expanding the determinant,

$$(1-\lambda)^2 - 9 = 0 \quad \text{or} \quad \lambda = -2, 4. \tag{3.6.10}$$

Thus, we have two real and distinct eigenvalues: $\lambda = -2$ and 4.

We must now find the corresponding $\mathbf{x}_0$ or *eigenvector* for each eigenvalue. From Equation 3.6.8,

$$(1-\lambda)a + 3b = 0, \tag{3.6.11}$$

and

$$3a + (1-\lambda)b = 0. \tag{3.6.12}$$

If $\lambda = 4$, these equations are consistent and yield $a = b = c_1$. If $\lambda = -2$, we have that $a = -b = c_2$. Therefore, the general solution in matrix notation is

$$\mathbf{x} = c_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{4t} + c_2\begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-2t}. \tag{3.6.13}$$

To evaluate c_1 and c_2, we must have initial conditions. For example, if $x_1(0) = x_2(0) = 1$, then

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{3.6.14}$$

Solving for c_1 and c_2, $c_1 = 1$, $c_2 = 0$, and the solution with this particular set of initial conditions is

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{4t}. \tag{3.6.15}$$

• Example 3.6.1

Let us solve the following set of linear ordinary differential equations

$$x_1' = -x_2 + x_3, \tag{3.6.16}$$

$$x_2' = 4x_1 - x_2 - 4x_3, \tag{3.6.17}$$

and

$$x_3' = -3x_1 - x_2 + 4x_3; \tag{3.6.18}$$

or in matrix form,

$$\mathbf{x}' = \begin{pmatrix} 0 & -1 & 1 \\ 4 & -1 & -4 \\ -3 & -1 & 4 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{3.6.19}$$

Assuming the solution $\mathbf{x} = \mathbf{x}_0 e^{\lambda t}$,

$$\begin{pmatrix} 0 & -1 & 1 \\ 4 & -1 & -4 \\ -3 & -1 & 4 \end{pmatrix} \mathbf{x}_0 = \lambda \mathbf{x}_0, \tag{3.6.20}$$

or

$$\begin{pmatrix} -\lambda & -1 & 1 \\ 4 & -1-\lambda & -4 \\ -3 & -1 & 4-\lambda \end{pmatrix} \mathbf{x}_0 = \mathbf{0}. \tag{3.6.21}$$

For nontrivial solutions,

$$\begin{vmatrix} -\lambda & -1 & 1 \\ 4 & -1-\lambda & -4 \\ -3 & -1 & 4-\lambda \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ 4-4\lambda & -5-\lambda & -4 \\ -3+4\lambda-\lambda^2 & 3-\lambda & 4-\lambda \end{vmatrix} = 0, \tag{3.6.22}$$

and

$$(\lambda-1)(\lambda-3)(\lambda+1) = 0, \quad \text{or} \quad \lambda = -1, 1, 3. \tag{3.6.23}$$

To determine the eigenvectors, we rewrite Equation 3.6.21 as

$$-\lambda a - b + c = 0, \tag{3.6.24}$$

$$4a - (1+\lambda)b - 4c = 0, \tag{3.6.25}$$

and

$$-3a - b + (4-\lambda)c = 0. \tag{3.6.26}$$

For example, if $\lambda = 1$,

$$-a - b + c = 0, \tag{3.6.27}$$

$$4a - 2b - 4c = 0, \tag{3.6.28}$$

and

$$-3a - b + 3c = 0; \tag{3.6.29}$$

or $a = c$, and $b = 0$. Thus, the eigenvector for $\lambda = 1$ is $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Similarly, for $\lambda = -1$, $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$; and for $\lambda = 3$, $\mathbf{x}_0 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$. Thus, the most general solution is

$$\mathbf{x} = c_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^t + c_2 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} e^{-t} + c_3 \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} e^{3t}. \tag{3.6.30}$$

□

• Example 3.6.2

Let us solve the following set of linear ordinary differential equations:

$$x_1' = x_1 - 2x_2, \tag{3.6.31}$$

and

$$x_2' = 2x_1 - 3x_2; \tag{3.6.32}$$

or in matrix form,

$$\mathbf{x}' = \begin{pmatrix} 1 & -2 \\ 2 & -3 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{3.6.33}$$

Assuming the solution $\mathbf{x} = \mathbf{x}_0 e^{\lambda t}$,

$$\begin{pmatrix} 1-\lambda & -2 \\ 2 & -3-\lambda \end{pmatrix} \mathbf{x}_0 = \mathbf{0}. \tag{3.6.34}$$

For nontrivial solutions,

$$\begin{vmatrix} 1-\lambda & -2 \\ 2 & -3-\lambda \end{vmatrix} = (\lambda+1)^2 = 0. \tag{3.6.35}$$

Thus, we have the solution

$$\mathbf{x} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t}. \tag{3.6.36}$$

The interesting aspect of this example is the single solution that the traditional approach yields because we have repeated roots. To find the second solution, we try the solution

$$\mathbf{x} = \begin{pmatrix} a+ct \\ b+dt \end{pmatrix} e^{-t}. \tag{3.6.37}$$

We guessed Equation 3.6.37 using our knowledge of solutions to differential equations when the characteristic polynomial has repeated roots. Substituting Equation 3.6.37 into Equation 3.6.33, we find that $c = d = 2c_2$, and $a - b = c_2$. Thus, we have one free parameter, which we choose to be b, and set it equal to zero. This is permissible because Equation 3.6.37 can be broken into two terms: $b \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t}$ and $c_2 \begin{pmatrix} 1+2t \\ 2t \end{pmatrix} e^{-t}$. The first term can be incorporated into the $c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t}$ term. Thus, the general solution is

$$\mathbf{x} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-t} + c_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-t} + 2c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} te^{-t}. \tag{3.6.38}$$

□

• Example 3.6.3

Let us solve the system of linear differential equations:

$$x_1' = 2x_1 - 3x_2, \tag{3.6.39}$$

and

$$x_2' = 3x_1 + 2x_2; \tag{3.6.40}$$

or in matrix form,

$$\mathbf{x}' = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{3.6.41}$$

Assuming the solution $\mathbf{x} = \mathbf{x}_0 e^{\lambda t}$,

$$\begin{pmatrix} 2-\lambda & -3 \\ 3 & 2-\lambda \end{pmatrix} \mathbf{x}_0 = \mathbf{0}. \tag{3.6.42}$$

For nontrivial solutions,

$$\begin{vmatrix} 2-\lambda & -3 \\ 3 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 + 9 = 0, \tag{3.6.43}$$

and $\lambda = 2 \pm 3i$. If $\mathbf{x}_0 = \begin{pmatrix} a \\ b \end{pmatrix}$, then $b = -ai$ if $\lambda = 2 + 3i$, and $b = ai$ if $\lambda = 2 - 3i$. Thus, the general solution is

$$\mathbf{x} = c_1 \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{2t+3it} + c_2 \begin{pmatrix} 1 \\ i \end{pmatrix} e^{2t-3it}, \tag{3.6.44}$$

where c_1 and c_2 are arbitrary complex constants. Using Euler relationships, we can rewrite Equation 3.6.44 as

$$\mathbf{x} = c_3 \begin{bmatrix} \cos(3t) \\ \sin(3t) \end{bmatrix} e^{2t} + c_4 \begin{bmatrix} \sin(3t) \\ -\cos(3t) \end{bmatrix} e^{2t}, \tag{3.6.45}$$

where $c_3 = c_1 + c_2$ and $c_4 = i(c_1 - c_2)$. □

• **Example 3.6.4: Diagonalization of a matrix A**

Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ denote the eigenvectors, with corresponding eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, of an $n \times n$ matrix A. If we introduce a matrix $X = [\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_n]$ (the eigenvectors form the columns of X) and recall that $A\mathbf{x}_j = \lambda_j \mathbf{x}_j$, then

$$AX = A[\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_n] = [A\mathbf{x}_1 A\mathbf{x}_2 \cdots A\mathbf{x}_n] = [\lambda_1 \mathbf{x}_1 \lambda_2 \mathbf{x}_2 \cdots \lambda_n \mathbf{x}_n]. \tag{3.6.46}$$

Therefore, $AX = XD$, where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \tag{3.6.47}$$

Because X has rank n, X^{-1} exists and $X^{-1}AX = X^{-1}XD = D$. Thus, $X^{-1}AX = D$ is a process whereby we can *diagonalize* the matrix A using the eigenvectors of A. Diagonalizable matrices are of interest because diagonal matrices are especially easy to use. Furthermore, we note that

$$D^2 = DD = X^{-1}AXX^{-1}AX = X^{-1}AAX = X^{-1}A^2X. \tag{3.6.48}$$

Repeating this process, we eventually obtain the general result that $D^m = X^{-1}A^mX$.

To verify $X^{-1}AX = D$, let us use

$$A = \begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix}. \tag{3.6.49}$$

This matrix has the eigenvalues $\lambda_{1,2} = 1, 5$ with the corresponding eigenvectors $\mathbf{x}_1 = (\,2 \quad 1\,)^T$ and $\mathbf{x}_2 = (\,2 \quad -1\,)^T$. Therefore,

$$X = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}, \qquad \text{and} \qquad X^{-1} = \begin{pmatrix} 1/4 & -1/2 \\ 1/4 & 1/2 \end{pmatrix}. \tag{3.6.50}$$

Therefore,

$$X^{-1}AX = \begin{pmatrix} 1/4 & -1/2 \\ 1/4 & 1/2 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} \tag{3.6.51}$$

$$= \begin{pmatrix} 1/4 & -1/2 \\ 1/4 & 1/2 \end{pmatrix} \begin{pmatrix} 2 & 10 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}. \tag{3.6.52}$$

Problems

Find the general solution of the following sets of ordinary differential equations using matrix technique. You may find the eigenvalues and eigenvectors either by hand or use MATLAB.

1. $x_1' = x_1 + 2x_2$ $\qquad x_2' = 2x_1 + x_2$

2. $x_1' = x_1 - 4x_2$ $\qquad x_2' = 3x_1 - 6x_2$

3. $x_1' = x_1 + x_2$ $\qquad x_2' = 4x_1 + x_2$

4. $x_1' = x_1 + 5x_2$ $\quad x_2' = -2x_1 - 6x_2$

5. $x_1' = -\frac{3}{2}x_1 - 2x_2$ $\quad x_2' = 2x_1 + \frac{5}{2}x_2.$

6. $x_1' = -3x_1 - 2x_2$ $\quad x_2' = 2x_1 + x_2$

7. $x_1' = x_1 - x_2$ $\quad x_2' = x_1 + 3x_2$

8. $x_1' = 3x_1 + 2x_2$ $\quad x_2' = -2x_1 - x_2$

9. $x_1' = -2x_1 - 13x_2$ $\quad x_2' = x_1 + 4x_2$

10. $x_1' = 3x_1 - 2x_2$ $\quad x_2' = 5x_1 - 3x_2$

11. $x_1' = 4x_1 - 2x_2$ $\quad x_2' = 25x_1 - 10x_2$

12. $x_1' = -3x_1 - 4x_2$ $\quad x_2' = 2x_1 + x_2$

13. $x_1' = 3x_1 + 4x_2$ $\quad x_2' = -2x_1 - x_2$

14. $x_1' + 5x_1 + x_2' + 3x_2 = 0$ $\quad 2x_1' + x_1 + x_2' + x_2 = 0$

15. $x_1' - x_1 + x_2' - 2x_2 = 0$ $\quad x_1' - 5x_1 + 2x_2' - 7x_2 = 0$

16. $x_1' = x_1 - 2x_2$ $\quad x_2' = 0$ $\quad x_3' = -5x_1 + 7x_3.$

17. $x_1' = 2x_1$ $\quad x_2' = x_1 + 2x_3$ $\quad x_3' = x_3.$

18. $x_1' = 3x_1 - 2x_3$ $\quad x_2' = -x_1 + 2x_2 + x_3$ $\quad x_3' = 4x_1 - 3x_3$

19. $x_1' = 3x_1 - x_3$ $\quad x_2' = -2x_1 + 2x_2 + x_3$ $\quad x_3' = 8x_1 - 3x_3$

3.7 MATRIX EXPONENTIAL

In the previous section we solved initial-value problems involving systems of linear ordinary differential equations via the eigenvalue problem. Here we introduce an alternative method based on the *matrix exponential*, defined by

$$e^{At} = I + At + \tfrac{1}{2!}A^2t^2 + \cdots + \tfrac{1}{k!}A^kt^k + \cdots. \tag{3.7.1}$$

Clearly

$$e^0 = e^{A0} = I, \qquad \text{and} \qquad \frac{d}{dt}\left(e^{At}\right) = Ae^{At}. \tag{3.7.2}$$

Therefore, using the matrix exponential function, the solution to the system of homogeneous linear first-order differential equations with constant coefficients

$$\mathbf{x}' = A\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{x}_0, \tag{3.7.3}$$

is $\mathbf{x}(t) = e^{At}\mathbf{x}_0$.

The question now arises as to how to compute this matrix exponential. There are several methods. For example, from the concept of diagonalization of a matrix (see Example 3.6.4) we can write $A = PDP^{-1}$, where P are the eigenvectors of A. Then,

$$e^A = \sum_{k=0}^{\infty} \frac{(PDP^{-1})^k}{k!} = \sum_{k=0}^{\infty} P \frac{D^k}{k!} P^{-1} = P \left(\sum_{k=0}^{\infty} \frac{D^k}{k!} \right) P^{-1} = Pe^D P^{-1}, \tag{3.7.4}$$

where

$$e^D = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix}. \tag{3.7.5}$$

Because many software packages contain routines for finding eigenvalues and eigenvectors, Equation 3.7.4 provides a convenient method for computing e^A. In the case of MATLAB, we just have to invoke the intrinsic function `expm(·)`.

In this section we focus on a recently developed method by Liz,[7] who improved a method constructed by Leonard.[8] The advantage of this method is that it uses techniques that we have already introduced. We will first state the result and then illustrate its use.

The main result of Liz's analysis is:

Theorem: *Let A be a constant $n \times n$ matrix with characteristic polynomial $p(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$. Then*

$$e^{At} = x_1(t)I + x_2(t)A + \cdots + x_n(t)A^{n-1}, \tag{3.7.6}$$

where

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} = B_0^{-1} \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix}, \tag{3.7.7}$$

$$B_t = \begin{pmatrix} \varphi_1(t) & \varphi_1'(t) & \cdots & \varphi_1^{(n-1)}(t) \\ \varphi_2(t) & \varphi_2'(t) & \cdots & \varphi_2^{(n-1)}(t) \\ \vdots & \vdots & \cdots & \vdots \\ \varphi_n(t) & \varphi_n'(t) & \cdots & \varphi_n^{(n-1)}(t) \end{pmatrix}, \tag{3.7.8}$$

and $S = \{\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)\}$ being a fundamental system of solutions for the homogeneous linear differential equations whose characteristic equation is the characteristic equation of A, $p(\lambda) = 0$. The proof is given in Liz's paper. Note that for this technique to work, $x_1(0) = 1$ and $x_2(0) = x_3(0) = \cdots = x_n(0) = 0$.

• Example 3.7.1

Let us illustrate this method of computing the matrix exponential by solving

$$x' = 2x - y + z, \tag{3.7.9}$$

[7] Liz, E., 1998: A note on the matrix exponential. *SIAM Rev.*, **40**, 700–702.

[8] Leonard, I. E., 1996: The matrix exponential. *SIAM Rev.*, **38**, 507–512.

$$y' = 3y - z, \tag{3.7.10}$$

and

$$z' = 2x + y + 3z. \tag{3.7.11}$$

The solution to this system of equations is $\mathbf{x} = e^{At}\mathbf{x}_0$, where

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}. \tag{3.7.12}$$

The vector $\mathbf{x}_0$ is the value of $\mathbf{x}(t)$ at $t = 0$.

Our first task is to compute the characteristic polynomial $p(\lambda) = 0$. This is simply

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 1 & -1 \\ 0 & \lambda - 3 & 1 \\ -2 & -1 & \lambda - 3 \end{vmatrix} = (\lambda - 2)^2(\lambda - 4) = 0. \tag{3.7.13}$$

Consequently, $\lambda = 2$ twice and $\lambda = 4$, and the fundamental solutions are $S = \{e^{2t}, te^{2t}, e^{4t}\}$. Therefore,

$$B_t = \begin{pmatrix} e^{4t} & 4e^{4t} & 16e^{4t} \\ e^{2t} & 2e^{2t} & 4e^{2t} \\ te^{2t} & e^{2t} + 2te^{2t} & 4e^{2t} + 4te^{2t} \end{pmatrix}, \tag{3.7.14}$$

and

$$A^2 = \begin{pmatrix} 6 & -4 & 6 \\ -2 & 8 & -6 \\ 10 & 4 & 10 \end{pmatrix}, \; B_0 = \begin{pmatrix} 1 & 4 & 16 \\ 1 & 2 & 4 \\ 0 & 1 & 4 \end{pmatrix}, \; B_0^{-1} = \begin{pmatrix} -1 & 0 & -4 \\ -1 & 1 & 3 \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} \end{pmatrix}. \tag{3.7.15}$$

The inverse B_0^{-1} can be found using either Gaussian elimination or MATLAB.

To find $x_1(t)$, $x_2(t)$ and $x_3(t)$, we have from Equation 3.7.7 that

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} -1 & 0 & -4 \\ -1 & 1 & 3 \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} e^{4t} \\ e^{2t} \\ te^{2t} \end{pmatrix} = \begin{pmatrix} e^{4t} - 4te^{2t} \\ e^{2t} - e^{4t} + 3te^{2t} \\ \frac{1}{4}e^{4t} - \frac{1}{4}e^{2t} - \frac{1}{2}te^{2t} \end{pmatrix}, \tag{3.7.16}$$

or

$$x_1(t) = e^{4t} - 4te^{2t}, \; x_2(t) = e^{2t} - e^{4t} + 3te^{2t}, \; x_3(t) = \tfrac{1}{4}e^{4t} - \tfrac{1}{4}e^{2t} - \tfrac{1}{2}te^{2t}. \tag{3.7.17}$$

Note that $x_1(0) = 1$ while $x_2(0) = x_3(0) = 0$.

Finally, we have that

$$e^{At} = x_1(t) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + x_2(t) \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{pmatrix} + x_3(t) \begin{pmatrix} 6 & -4 & 6 \\ -2 & 8 & -6 \\ 10 & 4 & 10 \end{pmatrix}. \tag{3.7.18}$$

Substituting for $x_1(t)$, $x_2(t)$ and $x_3(t)$ and simplifying, we finally obtain

$$e^{At} = \frac{1}{2} \begin{pmatrix} e^{4t} + e^{2t} - 2te^{2t} & -2te^{2t} & e^{4t} - e^{2t} \\ e^{2t} - e^{4t} + 2te^{2t} & 2(t+1)e^{2t} & e^{2t} - e^{4t} \\ e^{4t} - e^{2t} + 2te^{2t} & 2te^{2t} & e^{4t} + e^{2t} \end{pmatrix}; \tag{3.7.19}$$

or

$$x_1(t) = \left(\tfrac{1}{2}e^{4t} + \tfrac{1}{2}e^{2t} - te^{2t}\right) x_1(0) - te^{2t}x_2(0) + \left(\tfrac{1}{2}e^{4t} - \tfrac{1}{2}e^{2t}\right) x_3(0), \tag{3.7.20}$$

$$x_2(t) = \left(\tfrac{1}{2}e^{2t} - \tfrac{1}{2}e^{4t} + te^{2t}\right) x_1(0) + (t+1)e^{2t}x_2(0) + \left(\tfrac{1}{2}e^{2t} - \tfrac{1}{2}e^{4t}\right) x_3(0), \tag{3.7.21}$$

and

$$x_3(t) = \left(\tfrac{1}{2}e^{4t} - \tfrac{1}{2}e^{2t} + te^{2t}\right) x_1(0) + te^{2t}x_2(0) + \left(\tfrac{1}{2}e^{4t} + \tfrac{1}{2}e^{2t}\right) x_3(0). \tag{3.7.22}$$

□

• Example 3.7.2

The matrix exponential can also be used to solve systems of first-order, nonhomogeneous linear ordinary differential equations. To illustrate this, consider the following system of linear ordinary differential equations:

$$x' = x - 4y + e^{2t}, \tag{3.7.23}$$

and

$$y' = x + 5y + t. \tag{3.7.24}$$

We can rewrite this system as

$$\mathbf{x}' = A\mathbf{x} + \mathbf{b}, \tag{3.7.25}$$

where

$$A = \begin{pmatrix} 1 & -4 \\ 1 & 5 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} e^{2t} \\ t \end{pmatrix}, \qquad \text{and} \qquad \mathbf{x} = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}. \tag{3.7.26}$$

We leave as an exercise the computation of the matrix exponential and find that

$$e^{At} = \begin{pmatrix} e^{3t} - 2te^{3t} & -4te^{3t} \\ te^{3t} & e^{3t} + 2te^{3t} \end{pmatrix}. \tag{3.7.27}$$

Clearly the homogeneous solution is $\mathbf{x}_H(t) = e^{At}\mathbf{C}$, where $\mathbf{C}$ is the arbitrary constant that is determined by the initial condition. But how do we find the particular solution, $\mathbf{x}_p(t)$? Let $\mathbf{x}_p(t) = e^{At}\mathbf{y}(t)$. Then

$$\mathbf{x}_p'(t) = Ae^{At}\mathbf{y}(t) + e^{At}\mathbf{y}'(t), \tag{3.7.28}$$

or

$$\mathbf{x}_p'(t) = A\mathbf{x}_p(t) + e^{At}\mathbf{y}'(t). \tag{3.7.29}$$

Therefore,

$$e^{At}\mathbf{y}'(t) = \mathbf{b}(t), \qquad \text{or} \qquad \mathbf{y}(t) = e^{-At}\mathbf{b}(t), \tag{3.7.30}$$

since $\left(e^{A}\right)^{-1} = e^{-A}$. Integrating both sides of Equation 3.7.29 and multiplying through by e^{At}, we find that

$$\mathbf{x}_p(t) = \int_0^t e^{A(t-s)}\mathbf{b}(s)\,ds = \int_0^t e^{As}\mathbf{b}(t-s)\,ds. \tag{3.7.31}$$

Returning to our original problem,

$$\int_0^t e^{As}\mathbf{b}(t-s)\,ds = \int_0^t \begin{pmatrix} e^{3s}-2se^{3s} & -4se^{3s} \\ se^{3s} & e^{3s}+2se^{3s} \end{pmatrix} \begin{pmatrix} e^{2(t-s)} \\ t-s \end{pmatrix} ds \tag{3.7.32}$$

$$= \int_0^t \begin{pmatrix} e^{2t}\left(e^{s}-2se^{s}\right)-4s(t-s)e^{3s} \\ e^{2t}se^{s}+(t-s)e^{3s}+2s(t-s)e^{3s} \end{pmatrix} ds \tag{3.7.33}$$

$$= \begin{pmatrix} \frac{89}{27}e^{3t}-3e^{2t}-\frac{22}{9}te^{3t}-\frac{4}{9}t-\frac{8}{27} \\ \\ -\frac{28}{27}e^{3t}+e^{2t}+\frac{11}{9}te^{3t}-\frac{1}{9}t+\frac{1}{27} \end{pmatrix}. \tag{3.7.34}$$

The final answer consists of the homogeneous solution plus the particular solution.

Problems

Find e^{At} for the following matrices A:

1. $A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ 2. $A = \begin{pmatrix} 3 & 5 \\ 0 & 3 \end{pmatrix}$ 3. $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

4. $A = \begin{pmatrix} 2 & 3 & 4 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$ 5. $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$

For each of the following A's and $\mathbf{b}$'s, use the matrix exponential to find the general solution for the system of first-order, linear ordinary differential equations $\mathbf{x}' = A\mathbf{x} + \mathbf{b}$:

6. $A = \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} e^t \\ e^t \end{pmatrix}$ 7. $A = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} e^t \\ t \end{pmatrix}$

8. $A = \begin{pmatrix} 2 & 1 \\ -4 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} te^{2t} \\ -e^{2t} \end{pmatrix}$ 9. $A = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} -\cos(t) \\ \sin(t) \end{pmatrix}$

10. $A = \begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$ 11. $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ te^t \\ e^t \end{pmatrix}$

12. $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 2t \\ t+2 \\ 3t \end{pmatrix}$ 13. $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 0 \\ 4 & 0 & 1 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} -3e^t \\ 6e^t \\ -4e^t \end{pmatrix}$

14. $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} e^t \\ e^t \\ e^t \end{pmatrix}$ 15. $A = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 3 & 4 \\ 0 & 0 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} t \\ 1 \\ e^t \end{pmatrix}$.

Further Readings

Bronson, R., and G. B. Costa, 2007: *Linear Algebra: An Introduction.* Academic Press, 520 pp. Provides a step-by-step explanation of linear algebra.

Davis, H. T., and K. T. Thomson, 2000: *Linear Algebra and Linear Operators in Engineering with Applications in Mathematica.* Academic Press, 547 pp. Advanced textbook designed for first-year graduate students in the physical sciences and engineering.

Hoffman, J., 2001: *Numerical Methods for Engineers and Scientists.* Mc-Graw Hill, 823 pp. A first course in numerical methods that is both lucid and in depth.

Munakata, T., 1979: *Matrices and Linear Programming with Applications.* Holden-Day, 469 pp. Provides the basic concepts with clarity.

Noble, B., and J. W. Daniel, 1977: *Applied Linear Algebra.* Prentice Hall, 494 pp. Excellent conceptual explanations with well-motivated proofs.

Strang, G., 2009: *Linear Algebra and Its Applications.* Academic Press, 584 pp. A good supplement to a formal text that provides intuitive understanding.

Wilkinson, J. H., 1988: *The Algebraic Eigenvalue Problem.* Clarendon Press, 662 pp. The classic source book on the eigenvalue problem

Chapter 4
Vector Calculus

Physicists invented vectors and vector operations to facilitate their mathematical expression of such diverse topics as mechanics and electromagnetism. In this chapter we focus on multivariable differentiations and integrations of vector fields, such as the velocity of a fluid, where the vector field is solely a function of its position.

4.1 REVIEW

The physical sciences and engineering abound with vectors and scalars. *Scalars* are physical quantities that only possess magnitude. Examples include mass, temperature, density, and pressure. *Vectors* are physical quantities that possess both magnitude and direction. Examples include velocity, acceleration, and force. We shall denote vectors by boldfaced letters.

Two vectors are equal if they have the same magnitude and direction. From the limitless number of possible vectors, two special cases are the *zero vector* $\mathbf{0}$, which has no magnitude and unspecified direction, and the *unit vector*, which has unit magnitude.

The most convenient method for expressing a vector analytically is in terms of its components. A vector $\mathbf{a}$ in three-dimensional real space is any order triplet of real numbers (*components*) a_1, a_2, and a_3 such that $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, where $a_1\mathbf{i}$, $a_2\mathbf{j}$, and $a_3\mathbf{k}$ are vectors that lie along the coordinate axes and have their origin at a common initial point. The *magnitude*, *length*, or *norm* of a vector $\mathbf{a}$, $|\mathbf{a}|$, equals $\sqrt{a_1^2 + a_2^2 + a_3^2}$. A particularly important vector is the *position vector*, defined by $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

As in the case of scalars, certain arithmetic rules hold. Addition and subtraction are very similar to their scalar counterparts:

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k}, \qquad \textbf{(4.1.1)}$$

and

$$\mathbf{a} - \mathbf{b} = (a_1 - b_1)\mathbf{i} + (a_2 - b_2)\mathbf{j} + (a_3 - b_3)\mathbf{k}. \tag{4.1.2}$$

In contrast to its scalar counterpart, there are two types of multiplication. The *dot product* is defined as

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos(\theta) = a_1 b_1 + a_2 b_2 + a_3 b_3, \tag{4.1.3}$$

where θ is the angle between the vector such that $0 \le \theta \le \pi$. The dot product yields a scalar answer. A particularly important case is $\mathbf{a} \cdot \mathbf{b} = 0$ with $|\mathbf{a}| \neq 0$, and $|\mathbf{b}| \neq 0$. In this case the vectors are orthogonal (perpendicular) to each other.

The other form of multiplication is the *cross product*, which is defined by $\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}|\sin(\theta)\mathbf{n}$, where θ is the angle between the vectors such that $0 \le \theta \le \pi$, and $\mathbf{n}$ is a unit vector perpendicular to the plane of $\mathbf{a}$ and $\mathbf{b}$, with the direction given by the right-hand rule. A convenient method for computing the cross product from the scalar components of $\mathbf{a}$ and $\mathbf{b}$ is

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}. \tag{4.1.4}$$

Two nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ are *parallel* if and only if $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.

Most of the vectors that we will use are vector-valued functions. These functions are vectors that vary either with a single parametric variable t or multiple variables, say x, y, and z.

The most commonly encountered example of a vector-valued function that varies with a single independent variable involves the trajectory of particles. If a *space curve* is parameterized by the equations $x = f(t)$, $y = g(t)$, and $z = h(t)$ with $a \le t \le b$, the position vector $\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$ gives the location of a point P as it moves from its initial position to its final position. Furthermore, because the increment quotient $\Delta\mathbf{r}/\Delta t$ is in the direction of a secant line, then the limit of this quotient as $\Delta t \to 0$, $\mathbf{r}'(t)$ gives the tangent (tangent vector) to the curve at P.

• Example 4.1.1: Foucault pendulum

One of the great experiments of mid-nineteenth-century physics was the demonstration by J. B. L. Foucault (1819-1868) in 1851 of the earth's rotation by designing a (spherical) pendulum, supported by a long wire, that essentially swings in an nonaccelerating coordinate system. This problem demonstrates many of the fundamental concepts of vector calculus.

The total force[1] acting on the bob of the pendulum is $\mathbf{F} = \mathbf{T} + m\mathbf{G}$, where $\mathbf{T}$ is the tension in the pendulum and $\mathbf{G}$ is the gravitational attraction per unit mass. Using Newton's second law,

$$\left.\frac{d^2\mathbf{r}}{dt^2}\right|_{\text{inertial}} = \frac{\mathbf{T}}{m} + \mathbf{G}, \tag{4.1.5}$$

where $\mathbf{r}$ is the position vector from a fixed point in an inertial coordinate system to the bob. This system is inconvenient because we live on a rotating coordinate system. Employing

[1] See Broxmeyer, C., 1960: Foucault pendulum effect in a Schuler-tuned system. *J. Aerosp. Sci.*, **27**, 343–347.

the conventional geographic coordinate system,[2] Equation 4.1.5 becomes

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) = \frac{\mathbf{T}}{m} + \mathbf{G}, \tag{4.1.6}$$

where $\boldsymbol{\Omega}$ is the angular rotation vector of the earth and $\mathbf{r}$ now denotes a position vector in the rotating reference system with its origin at the center of the earth and terminal point at the bob. If we define the gravity vector $\mathbf{g} = \mathbf{G} - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r})$, then the dynamical equation is

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} = \frac{\mathbf{T}}{m} + \mathbf{g}, \tag{4.1.7}$$

where the second term on the left side of Equation 4.1.7 is called the *Coriolis force*.

Because the equation is *linear*, let us break the position vector $\mathbf{r}$ into two separate vectors: $\mathbf{r}_0$ and $\mathbf{r}_1$, where $\mathbf{r} = \mathbf{r}_0 + \mathbf{r}_1$. The vector $\mathbf{r}_0$ extends from the center of the earth to the pendulum's point of support, and $\mathbf{r}_1$ extends from the support point to the bob. Because $\mathbf{r}_0$ is a constant in the geographic system,

$$\frac{d^2\mathbf{r}_1}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}_1}{dt} = \frac{\mathbf{T}}{m} + \mathbf{g}. \tag{4.1.8}$$

If the length of the pendulum is L, then for small oscillations $\mathbf{r}_1 \approx x\mathbf{i} + y\mathbf{j} + L\mathbf{k}$ and the equations of motion are

$$\frac{d^2x}{dt^2} + 2\Omega\sin(\lambda)\frac{dy}{dt} = \frac{T_x}{m}, \tag{4.1.9}$$

$$\frac{d^2y}{dt^2} - 2\Omega\sin(\lambda)\frac{dx}{dt} = \frac{T_y}{m}, \tag{4.1.10}$$

and

$$2\Omega\cos(\lambda)\frac{dy}{dt} - g = \frac{T_z}{m}, \tag{4.1.11}$$

where λ denotes the latitude of the point and Ω is the rotation rate of the earth. The relationships between the components of tension are $T_x = xT_z/L$, and $T_y = yT_z/L$. From Equation 4.1.11,

$$\frac{T_z}{m} + g = 2\Omega\cos(\lambda)\frac{dy}{dt} \approx 0. \tag{4.1.12}$$

Substituting the definitions of T_x, T_y, and Equation 4.1.12 into Equation 4.1.9 and Equation 4.1.10,

$$\frac{d^2x}{dt^2} + \frac{g}{L}x + 2\Omega\sin(\lambda)\frac{dy}{dt} = 0, \tag{4.1.13}$$

and

$$\frac{d^2y}{dt^2} + \frac{g}{L}y - 2\Omega\sin(\lambda)\frac{dx}{dt} = 0. \tag{4.1.14}$$

The approximate solution to these coupled differential equations is

$$x(t) \approx A_0\cos[\Omega\sin(\lambda)t]\,\sin\left(\sqrt{g/L}\;t\right), \tag{4.1.15}$$

and

$$y(t) \approx A_0\sin[\Omega\sin(\lambda)t]\,\sin\left(\sqrt{g/L}\;t\right), \tag{4.1.16}$$

[2] For the derivation, see Marion, J. B., 1965: *Classical Dynamics of Particles and Systems*. Academic Press, Sections 12.2–12.3.

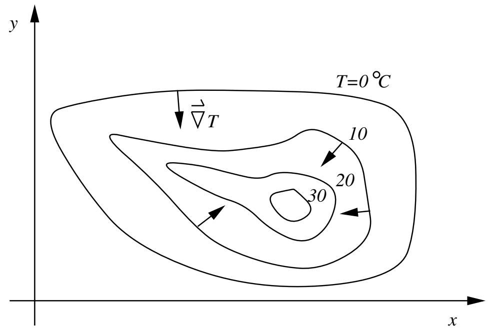


Figure 4.1.1: For a two-dimensional field $T(x, y)$, the gradient is a vector that is perpendicular to the isotherms $T(x, y) =$ constant and points in the direction of most rapidly increasing temperatures.

if $\Omega^2 \ll g/L$. Thus, we have a pendulum that swings with an angular frequency $\sqrt{g/L}$. However, depending upon the *latitude* λ, the direction in which the pendulum swings changes counterclockwise with time, completing a full cycle in $2\pi/[\Omega\sin(\lambda)]$. This result is most clearly seen when $\lambda = \pi/2$ and we are at the North Pole. There the earth is turning underneath the pendulum. If initially we set the pendulum swinging along the 0° longitude, the pendulum will shift with time to longitudes east of the Greenwich median. Eventually, after 24 hours, the process repeats itself. □

Consider now vector-valued functions that vary with several variables. A *vector function of position* assigns a vector value for every value of x, y, and z within some domain. Examples include the velocity field of a fluid at a given instant:

$$\mathbf{v} = u(x, y, z)\mathbf{i} + v(x, y, z)\mathbf{j} + w(x, y, z)\mathbf{k}. \qquad (4.1.17)$$

Another example arises in electromagnetism where electric and magnetic fields often vary as a function of the space coordinates. For us, however, probably the most useful example involves the vector differential operator, *del* or *nabla*,

$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}, \qquad (4.1.18)$$

which we apply to the multivariable differentiable scalar function $F(x, y, z)$ to give the *gradient* ∇F.

An important geometric interpretation of the gradient-one which we shall use frequently-is the fact that ∇f is perpendicular (normal) to the level surface at a given point P. To prove this, let the equation $F(x, y, z) = c$ describe a three-dimensional surface. If the differentiable functions $x = f(t)$, $y = g(t)$, and $z = h(t)$ are the parametric equations of a curve on the surface, then the derivative of $F[f(t), g(t), h(t)] = c$ is

$$\frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt} = 0, \qquad (4.1.19)$$

or

$$\nabla F \cdot \mathbf{r}' = 0. \qquad (4.1.20)$$

When $\mathbf{r}' \neq \mathbf{0}$, the vector ∇F is orthogonal to the tangent vector. Because our argument holds for any differentiable curve that passes through the arbitrary point (x, y, z), then ∇F is normal to the level surface at that point.

Figure 4.1.1 gives a common application of the gradient. Consider a two-dimensional temperature field $T(x, y)$. The level curves $T(x, y)$ = constant are lines that connect points where the temperature is the same (isotherms). The gradient in this case ∇T is a vector that is perpendicular or normal to these isotherms and points in the direction of most rapidly increasing temperature.

• Example 4.1.2

Let us find the gradient of the function $f(x, y, z) = x^2 z^2 \sin(4y)$.
Using the definition of gradient,

$$\nabla f = \frac{\partial[x^2 z^2 \sin(4y)]}{\partial x}\mathbf{i} + \frac{\partial[x^2 z^2 \sin(4y)]}{\partial y}\mathbf{j} + \frac{\partial[x^2 z^2 \sin(4y)]}{\partial z}\mathbf{k} \qquad (\mathbf{4.1.21})$$

$$= 2xz^2 \sin(4y)\mathbf{i} + 4x^2 z^2 \cos(4y)\mathbf{j} + 2x^2 z \sin(4y)\mathbf{k}. \qquad (\mathbf{4.1.22})$$

□

• Example 4.1.3

Let us find the unit normal to the unit sphere at any arbitrary point (x, y, z).
The surface of a unit sphere is defined by the equation $f(x, y, z) = x^2 + y^2 + z^2 = 1$. Therefore, the normal is given by the gradient

$$\mathbf{N} = \nabla f = 2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}, \qquad (\mathbf{4.1.23})$$

and the unit normal

$$\mathbf{n} = \frac{\nabla f}{|\nabla f|} = \frac{2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}}{\sqrt{4x^2 + 4y^2 + 4z^2}} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \qquad (\mathbf{4.1.24})$$

because $x^2 + y^2 + z^2 = 1$. □

• Example 4.1.4

In Figure 4.1.2, MATLAB has been used to illustrate the unit normal of the surface $z = 4 - x^2 - y^2$. Here $f(x, y, z) = z + x^2 + y^2 = 4$ so that $\nabla f = 2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k}$. The corresponding script is:

```
clear % clear variables
clf % clear figures
[x,y] = meshgrid(-2:0.5:2); % create the grid
z = 4 - x.^2 - y.^2; % compute surface within domain
% compute the gradient of f(x,y,z) = z + x^2 + y^2 = 4
% the x, y, and z components are u, v, and w
u = 2*x; v = 2*y; w = 1;
% find magnitude of gradient at each point
magnitude = sqrt(u.*u + v.*v + w.*w);
% compute unit gradient vector
u = u./magnitude; v = v./magnitude; w = w./magnitude;
mesh(x,y,z) % plot the surface
axis square
xlabel('x'); ylabel('y')
```

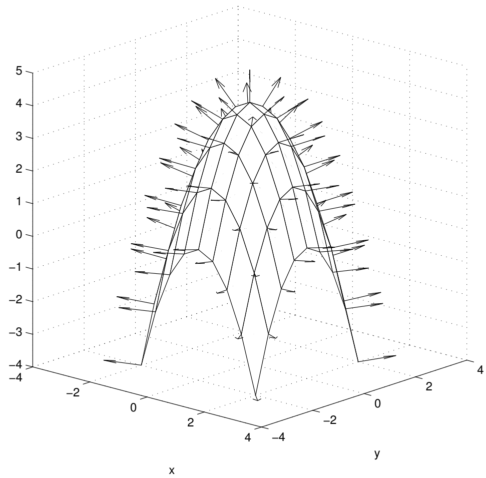


Figure 4.1.2: MATLAB plot of the function $z = 4 - x^2 - y^2$. The arrows give the unit normal to this surface.

```
hold on
% plot the unit gradient vector
quiver3(x,y,z,u,v,w,0)
```

This figure clearly shows that gradient gives a vector which is perpendicular to the surface. □

A popular method for visualizing a vector field $\mathbf{F}$ is to draw space curves that are tangent to the vector field at each x, y, z. In fluid mechanics these lines are called *streamlines* while in physics they are generally called *lines of force* or *flux lines* for an electric, magnetic, or gravitational field. For a fluid with a velocity field that does not vary with time, the streamlines give the paths along which small parcels of the fluid move.

To find the streamlines of a given vector field $\mathbf{F}$ with components $P(x, y, z)$, $Q(x, y, z)$, and $R(x, y, z)$, we assume that we can parameterize the streamlines in the form $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$. Then the tangent line is $\mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}$. Because the streamline must be parallel to the vector field at any t, $\mathbf{r}'(t) = \lambda\mathbf{F}$, or

$$\frac{dx}{dt} = \lambda P(x, y, z), \quad \frac{dy}{dt} = \lambda Q(x, y, z), \quad \text{and} \quad \frac{dz}{dt} = \lambda R(x, y, z), \tag{4.1.25}$$

or

$$\frac{dx}{P(x, y, z)} = \frac{dy}{Q(x, y, z)} = \frac{dz}{R(x, y, z)}. \tag{4.1.26}$$

The solution of this system of differential equations yields the streamlines.

• Example 4.1.5

Let us find the streamlines for the vector field $\mathbf{F} = \sec(x)\mathbf{i} - \cot(y)\mathbf{j} + \mathbf{k}$ that passes through the point $(\pi/4, \pi, 1)$. In this particular example, $\mathbf{F}$ represents a measured or computed fluid's velocity at a particular instant.

From Equation 4.1.26,

$$\frac{dx}{\sec(x)} = -\frac{dy}{\cot(y)} = \frac{dz}{1}. \tag{4.1.27}$$

This yields two differential equations:

$$\cos(x)\,dx = -\frac{\sin(y)}{\cos(y)}\,dy, \quad \text{and} \quad dz = -\frac{\sin(y)}{\cos(y)}\,dy. \tag{4.1.28}$$

Integrating these equations gives

$$\sin(x) = \ln|\cos(y)| + c_1, \quad \text{and} \quad z = \ln|\cos(y)| + c_2. \tag{4.1.29}$$

Substituting for the given point, we finally have that

$$\sin(x) = \ln|\cos(y)| + \sqrt{2}/2, \quad \text{and} \quad z = \ln|\cos(y)| + 1. \tag{4.1.30}$$

□

• Example 4.1.6

Let us find the streamlines for the vector field $\mathbf{F} = \sin(z)\mathbf{j} + e^y\mathbf{k}$ that passes through the point $(2, 0, 0)$.

From Equation 4.1.26,

$$\frac{dx}{0} = \frac{dy}{\sin(z)} = \frac{dz}{e^y}. \tag{4.1.31}$$

This yields two differential equations:

$$dx = 0, \qquad \text{and} \qquad \sin(z)\,dz = e^y\,dy. \tag{4.1.32}$$

Integrating these equations gives

$$x = c_1, \qquad \text{and} \qquad e^y = -\cos(z) + c_2. \tag{4.1.33}$$

Substituting for the given point, we finally have that

$$x = 2, \qquad \text{and} \qquad e^y = 2 - \cos(z). \tag{4.1.34}$$

Note that Equation 4.1.34 only applies for a certain strip in the yz-plane.

Problems

Given the following vectors $\mathbf{a}$ and $\mathbf{b}$, verify that $\mathbf{a}\cdot(\mathbf{a}\times\mathbf{b}) = 0$, and $\mathbf{b}\cdot(\mathbf{a}\times\mathbf{b}) = 0$:

1. $\mathbf{a} = 4\mathbf{i} - 2\mathbf{j} + 5\mathbf{k}, \quad \mathbf{b} = 3\mathbf{i} + \mathbf{j} - \mathbf{k}$
2. $\mathbf{a} = \mathbf{i} - 3\mathbf{j} + \mathbf{k}, \quad \mathbf{b} = 2\mathbf{i} + 4\mathbf{k}$
3. $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}, \quad \mathbf{b} = -5\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$
4. $\mathbf{a} = 8\mathbf{i} + \mathbf{j} - 6\mathbf{k}, \quad \mathbf{b} = \mathbf{i} - 2\mathbf{j} + 10\mathbf{k}$
5. $\mathbf{a} = 2\mathbf{i} + 7\mathbf{j} - 4\mathbf{k}, \quad \mathbf{b} = \mathbf{i} + \mathbf{j} - \mathbf{k}$.
6. Prove $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$.
7. Prove $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) + \mathbf{b}\times(\mathbf{c}\times\mathbf{a}) + \mathbf{c}\times(\mathbf{a}\times\mathbf{b}) = \mathbf{0}$.

Find the gradient of the following functions:

8. $f(x,y,z) = xy^2/z^3$ 9. $f(x,y,z) = xy\cos(yz)$ 10. $f(x,y,z) = \ln(x^2+y^2+z^2)$

11. $f(x,y,z) = x^2y^2(2z+1)^2$ 12. $f(x,y,z) = 2x - y^2 + z^2$.

Use MATLAB to illustrate the following surfaces as well as the unit normal.

13. $z = 3$ 14. $x^2 + y^2 = 4$ 15. $z = x^2 + y^2$ 16. $z = \sqrt{x^2+y^2}$

17. $z = y$ 18. $x + y + z = 1$ 19. $z = x^2$.

Find the streamlines for the following vector fields that pass through the specified point:

20. $\mathbf{F} = \mathbf{i} + \mathbf{j} + \mathbf{k}$; $(0,1,1)$ 21. $\mathbf{F} = 2\mathbf{i} - y^2\mathbf{j} + z\mathbf{k}$; $(1,1,1)$

22. $\mathbf{F} = 3x^2\mathbf{i} - y^2\mathbf{j} + z^2\mathbf{k}$; $(2,1,3)$ 23. $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} - z^3\mathbf{k}$; $(1,1,1)$

24. $\mathbf{F} = (1/x)\mathbf{i} + e^y\mathbf{j} - \mathbf{k}$; $(2,0,4)$.

25. Solve the differential equations, Equation 4.1.13 and Equation 4.1.14 with the initial conditions $x(0) = y(0) = y'(0) = 0$, and $x'(0) = A_0\sqrt{g/L}$ assuming that $\Omega^2 \ll g/L$.

26. If a fluid is bounded by a fixed surface $f(x,y,z) = c$, show that the fluid must satisfy the boundary condition $\mathbf{v} \cdot \nabla f = 0$, where $\mathbf{v}$ is the velocity of the fluid.

27. A sphere of radius a is moving in a fluid with the constant velocity $\mathbf{u}$. Show that the fluid satisfies the boundary condition $(\mathbf{v} - \mathbf{u}) \cdot (\mathbf{r} - \mathbf{u}t) = 0$ at the surface of the sphere, if the center of the sphere coincides with the origin at $t = 0$ and $\mathbf{v}$ denotes the velocity of the fluid.

4.2 DIVERGENCE AND CURL

Consider a vector field $\mathbf{v}$ defined in some region of three-dimensional space. The function $\mathbf{v}(\mathbf{r})$ can be resolved into components along the $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ directions, or

$$\mathbf{v}(\mathbf{r}) = u(x,y,z)\mathbf{i} + v(x,y,z)\mathbf{j} + w(x,y,z)\mathbf{k}. \quad \mathbf{(4.2.1)}$$

If $\mathbf{v}$ is a fluid's velocity field, then we can compute the flow rate through a small (differential) rectangular box defined by increments $(\Delta x, \Delta y, \Delta z)$ centered at the point (x,y,z). See Figure 4.2.1. The flow out from the box through the face with the outwardly pointing normal $\mathbf{n} = -\mathbf{j}$ is

$$\mathbf{v} \cdot (-\mathbf{j}) = -v(x, y - \Delta y/2, z)\Delta x \Delta z, \quad \mathbf{(4.2.2)}$$

and the flow through the face with the outwardly pointing normal $\mathbf{n} = \mathbf{j}$ is

$$\mathbf{v} \cdot \mathbf{j} = v(x, y + \Delta y/2, z)\Delta x \Delta z. \quad \mathbf{(4.2.3)}$$

The net flow through the two faces is

$$[v(x, y + \Delta y/2, z) - v(x, y - \Delta y/2, z)]\Delta x \Delta z \approx v_y(x,y,z)\Delta x \Delta y \Delta z. \quad \mathbf{(4.2.4)}$$

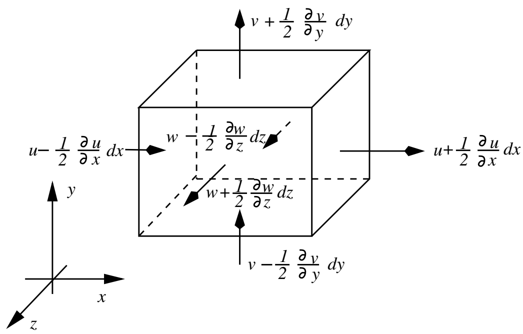


Figure 4.2.1: Divergence of a vector function $\mathbf{v}(x, y, z)$.

A similar analysis of the other faces and combination of the results give the approximate total flow from the box as

$$[u_x(x, y, z) + v_y(x, y, z) + w_z(x, y, z)]\Delta x\Delta y\Delta z. \tag{4.2.5}$$

Dividing by the volume $\Delta x\Delta y\Delta z$ and taking the limit as the dimensions of the box tend to zero yield $u_x + v_y + w_z$ as the flow out from (x, y, z) per unit volume per unit time. This scalar quantity is called the *divergence* of the vector $\mathbf{v}$:

$$\text{div}(\mathbf{v}) = \nabla \cdot \mathbf{v} = \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) \cdot (u\mathbf{i} + v\mathbf{j} + w\mathbf{k}) = u_x + v_y + w_z. \tag{4.2.6}$$

Thus, if the divergence is positive, either the fluid is expanding and its density at the point is falling with time, or the point is a *source* at which fluid is entering the field. When the divergence is negative, either the fluid is contracting and its density is rising at the point, or the point is a negative source or *sink* at which fluid is leaving the field.

If the divergence of a vector field is zero everywhere within a domain, then the flux entering any element of space exactly balances that leaving it and the vector field is called *nondivergent* or *solenoidal* (from a Greek word meaning a tube). For a fluid, if there are no sources or sinks, then its density cannot change.

Some useful properties of the divergence operator are

$$\nabla \cdot (\mathbf{F} + \mathbf{G}) = \nabla \cdot \mathbf{F} + \nabla \cdot \mathbf{G}, \tag{4.2.7}$$

$$\nabla \cdot (\varphi\mathbf{F}) = \varphi\nabla \cdot \mathbf{F} + \mathbf{F} \cdot \nabla\varphi \tag{4.2.8}$$

and

$$\nabla^2\varphi = \nabla \cdot \nabla\varphi = \varphi_{xx} + \varphi_{yy} + \varphi_{zz}. \tag{4.2.9}$$

Equation 4.2.9 is very important in physics and is given the special name of the *Laplacian.*[3]

• Example 4.2.1

If $\mathbf{F} = x^2z\mathbf{i} - 2y^3z^2\mathbf{j} + xy^2z\mathbf{k}$, compute the divergence of $\mathbf{F}$.

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}\left(x^2z\right) + \frac{\partial}{\partial y}\left(-2y^3z^2\right) + \frac{\partial}{\partial z}\left(xy^2z\right) = 2xz - 6y^2z^2 + xy^2. \tag{4.2.10}$$

[3] Some mathematicians write Δ instead of ∇^2.

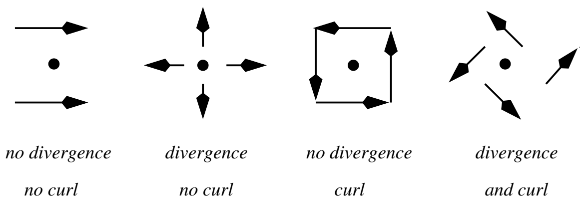


Figure 4.2.2: Examples of vector fields with and without divergence and curl.

□

• Example 4.2.2

If $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, show that $\mathbf{r}/|\mathbf{r}|^3$ is nondivergent.

$$\nabla \cdot \left(\frac{\mathbf{r}}{|\mathbf{r}|^3} \right) = \frac{\partial}{\partial x}\left[\frac{x}{(x^2+y^2+z^2)^{3/2}} \right] + \frac{\partial}{\partial y}\left[\frac{y}{(x^2+y^2+z^2)^{3/2}} \right] + \frac{\partial}{\partial z}\left[\frac{z}{(x^2+y^2+z^2)^{3/2}} \right] \tag{4.2.11}$$

$$= \frac{3}{(x^2+y^2+z^2)^{3/2}} - \frac{3x^2+3y^2+3z^2}{(x^2+y^2+z^2)^{5/2}} = 0. \tag{4.2.12}$$

□

Another important vector function involving the vector field $\mathbf{v}$ is the curl of $\mathbf{v}$, written curl($\mathbf{v}$) or rot($\mathbf{v}$) in some older textbooks. In fluid flow problems it is proportional to the instantaneous angular velocity of a fluid element. In rectangular coordinates,

$$\text{curl}(\mathbf{v}) = \nabla \times \mathbf{v} = (w_y - v_z)\mathbf{i} + (u_z - w_x)\mathbf{j} + (v_x - u_y)\mathbf{k}, \tag{4.2.13}$$

where $\mathbf{v} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$ as before. However, it is best remembered in the mnemonic form:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u & v & w \end{vmatrix} = (w_y - v_z)\mathbf{i} + (u_z - w_x)\mathbf{j} + (v_x - u_y)\mathbf{k}. \tag{4.2.14}$$

If the curl of a vector field is zero everywhere within a region, then the field is *irrotational.*

Figure 4.2.2 illustrates graphically some vector fields that do and do not possess divergence and curl. Let the vectors that are illustrated represent the motion of fluid particles. In the case of divergence only, fluid is streaming from the point, at which the density is falling. Alternatively the point could be a source. In the case where there is only curl, the fluid rotates about the point and the fluid is incompressible. Finally, the point that possesses both divergence and curl is a compressible fluid with rotation.

Some useful computational formulas exist for both the divergence and curl operations:

$$\nabla \times (\mathbf{F} + \mathbf{G}) = \nabla \times \mathbf{F} + \nabla \times \mathbf{G}, \tag{4.2.15}$$

$$\nabla \times \nabla\varphi = \mathbf{0}, \tag{4.2.16}$$

$$\nabla \cdot \nabla \times \mathbf{F} = 0, \tag{4.2.17}$$

$$\nabla \times (\varphi \mathbf{F}) = \varphi \nabla \times \mathbf{F} + \nabla \varphi \times \mathbf{F}, \tag{4.2.18}$$

$$\nabla(\mathbf{F} \cdot \mathbf{G}) = (\mathbf{F} \cdot \nabla)\mathbf{G} + (\mathbf{G} \cdot \nabla)\mathbf{F} + \mathbf{F} \times (\nabla \times \mathbf{G}) + \mathbf{G} \times (\nabla \times \mathbf{F}), \tag{4.2.19}$$

$$\nabla \times (\mathbf{F} \times \mathbf{G}) = (\mathbf{G} \cdot \nabla)\mathbf{F} - (\mathbf{F} \cdot \nabla)\mathbf{G} + \mathbf{F}(\nabla \cdot \mathbf{G}) - \mathbf{G}(\nabla \cdot \mathbf{F}), \tag{4.2.20}$$

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - (\nabla \cdot \nabla)\mathbf{F}, \tag{4.2.21}$$

and

$$\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot \nabla \times \mathbf{F} - \mathbf{F} \cdot \nabla \times \mathbf{G}. \tag{4.2.22}$$

In this book the operation $\nabla \mathbf{F}$ is undefined.

• **Example 4.2.3**

If $\mathbf{F} = xz^3\mathbf{i} - 2x^2yz\mathbf{j} + 2yz^4\mathbf{k}$, compute the curl of $\mathbf{F}$ and verify that $\nabla \cdot \nabla \times \mathbf{F} = 0$. From the definition of curl,

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xz^3 & -2x^2yz & 2yz^4 \end{vmatrix} \tag{4.2.23}$$

$$= \left[\frac{\partial}{\partial y}\left(2yz^4\right) - \frac{\partial}{\partial z}\left(-2x^2yz\right)\right]\mathbf{i} - \left[\frac{\partial}{\partial x}\left(2yz^4\right) - \frac{\partial}{\partial z}\left(xz^3\right)\right]\mathbf{j}$$
$$+ \left[\frac{\partial}{\partial x}\left(-2x^2yz\right) - \frac{\partial}{\partial y}\left(xz^3\right)\right]\mathbf{k} \tag{4.2.24}$$

$$= (2z^4 + 2x^2y)\mathbf{i} - (0 - 3xz^2)\mathbf{j} + (-4xyz - 0)\mathbf{k} \tag{4.2.25}$$

$$= (2z^4 + 2x^2y)\mathbf{i} + 3xz^2\mathbf{j} - 4xyz\mathbf{k}. \tag{4.2.26}$$

From the definition of divergence and Equation 4.2.26,

$$\nabla \cdot \nabla \times \mathbf{F} = \frac{\partial}{\partial x}\left(2z^4 + 2x^2y\right) + \frac{\partial}{\partial y}\left(3xz^2\right) + \frac{\partial}{\partial z}\left(-4xyz\right) = 4xy + 0 - 4xy = 0. \tag{4.2.27}$$

□

• **Example 4.2.4: Potential flow theory**

One of the topics in most elementary fluid mechanics courses is the study of irrotational and nondivergent fluid flows. Because the fluid is irrotational, the velocity vector field $\mathbf{v}$ satisfies $\nabla \times \mathbf{v} = \mathbf{0}$. From Equation 4.2.16 we can introduce a potential φ such that $\mathbf{v} = \nabla\varphi$. Because the flow field is nondivergent, $\nabla \cdot \mathbf{v} = \nabla^2\varphi = 0$. Thus, the fluid flow can be completely described in terms of solutions to Laplace's equation. This area of fluid mechanics is called *potential flow theory.*

Problems

Compute $\nabla \cdot \mathbf{F}, \nabla \times \mathbf{F}, \nabla \cdot (\nabla \times \mathbf{F})$, and $\nabla(\nabla \cdot \mathbf{F})$, for the following vector fields:

1. $\mathbf{F} = x^2z\mathbf{i} + yz^2\mathbf{j} + xy^2\mathbf{k}$
2. $\mathbf{F} = 4x^2y^2\mathbf{i} + (2x + 2yz)\mathbf{j} + (3z + y^2)\mathbf{k}$
3. $\mathbf{F} = (x - y)^2\mathbf{i} + e^{-xy}\mathbf{j} + xze^{2y}\mathbf{k}$
4. $\mathbf{F} = 3xy\mathbf{i} + 2xz^2\mathbf{j} + y^3\mathbf{k}$
5. $\mathbf{F} = 5yz\mathbf{i} + x^2z\mathbf{j} + 3x^3\mathbf{k}$
6. $\mathbf{F} = y^3\mathbf{i} + (x^3y^2 - xy)\mathbf{j} - (x^3yz - xz)\mathbf{k}$

7. $\mathbf{F} = xe^{-y}\mathbf{i} + yz^2\mathbf{j} + 3e^{-z}\mathbf{k}$

8. $\mathbf{F} = y\ln(x)\mathbf{i} + (2 - 3yz)\mathbf{j} + xyz^3\mathbf{k}$

9. $\mathbf{F} = xyz\mathbf{i} + x^3yze^z\mathbf{j} + xye^z\mathbf{k}$

10. $\mathbf{F} = (xy^3 - z^4)\mathbf{i} + 4x^4y^2z\mathbf{j} - y^4z^5\mathbf{k}$

11. $\mathbf{F} = xy^2\mathbf{i} + xyz^2\mathbf{j} + xy\cos(z)\mathbf{k}$

12. $\mathbf{F} = xy^2\mathbf{i} + xyz^2\mathbf{j} + xy\sin(z)\mathbf{k}$

13. $\mathbf{F} = xy^2\mathbf{i} + xyz\mathbf{j} + xy\cos(z)\mathbf{k}$

14. (a) Assuming continuity of all partial derivatives, show that

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2\mathbf{F}.$$

(b) Using $\mathbf{F} = 3xy\mathbf{i} + 4yz\mathbf{j} + 2xz\mathbf{k}$, verify the results in part (a).

15. If $\mathbf{E} = \mathbf{E}(x, y, z, t)$ and $\mathbf{B} = \mathbf{B}(x, y, z, t)$ represent the electric and magnetic fields in a vacuum, Maxwell's field equations are:

$$\nabla \cdot \mathbf{E} = 0, \qquad \nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t},$$

$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t},$$

where c is the speed of light. Using the results from Problem 14, show that $\mathbf{E}$ and $\mathbf{B}$ satisfy

$$\nabla^2\mathbf{E} = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad \text{and} \quad \nabla^2\mathbf{B} = \frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2}.$$

16. If f and g are continuously differentiable scalar fields, show that $\nabla f \times \nabla g$ is solenoidal. Hint: Show that $\nabla f \times \nabla g = \nabla \times (f\nabla g)$.

17. An inviscid (frictionless) fluid in equilibrium obeys the relationship $\nabla p = \rho\mathbf{F}$, where ρ denotes the density of the fluid, p denotes the pressure, and $\mathbf{F}$ denotes the body forces (such as gravity). Show that $\mathbf{F} \cdot \nabla \times \mathbf{F} = 0$.

4.3 LINE INTEGRALS

Line integrals are ubiquitous in physics. In mechanics they are used to compute work. In electricity and magnetism, they provide simple methods for computing the electric and magnetic fields for simple geometries.

The line integral most frequently encountered is an *oriented* one in which the path C is directed and the integrand is the dot product between the vector function $\mathbf{F}(\mathbf{r})$ and the tangent of the path $d\mathbf{r}$. It is usually written in the economical form

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz, \tag{4.3.1}$$

where $\mathbf{F} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$. If the starting and terminal points are the same so that the contour is closed, then this *closed contour integral* will be denoted by $\oint_C$.

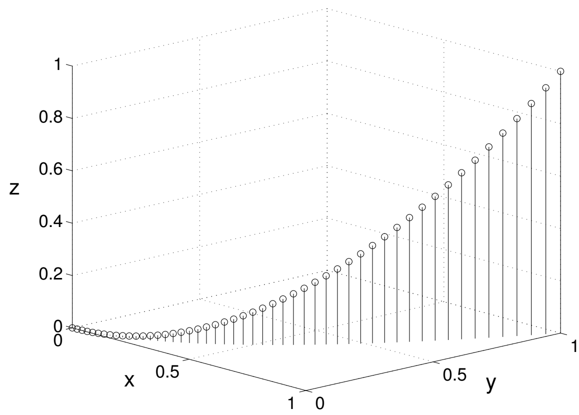


Figure 4.3.1: Diagram for the line integration in Example 4.3.1.

In the following examples we show how to evaluate the line integrals along various types of curves.

• Example 4.3.1

If $\mathbf{F} = (3x^2 + 6y)\mathbf{i} - 14yz\mathbf{j} + 20xz^2\mathbf{k}$, let us evaluate the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ along the parametric curves $x(t) = t$, $y(t) = t^2$, and $z(t) = t^3$ from the point $(0, 0, 0)$ to $(1, 1, 1)$. Using the MATLAB commands

```
>> clear
>> t = 0:0.02:1
>> stem3(t,t.^2,t.^3); xlabel('x','Fontsize',20); ...
   ylabel('y','Fontsize',20); zlabel('z','Fontsize',20);
```

we illustrate these parametric curves in Figure 4.3.1.

We begin by finding the values of t, which give the corresponding endpoints. A quick check shows that $t = 0$ gives $(0, 0, 0)$ while $t = 1$ yields $(1, 1, 1)$. It should be noted that the same value of t must give the correct coordinates in each direction. Failure to do so suggests an error in the parameterization. Therefore,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (3t^2 + 6t^2)\, dt - 14t^2(t^3)\, d(t^2) + 20t(t^3)^2 d(t^3) \tag{4.3.2}$$

$$= \int_0^1 9t^2\, dt - 28t^6\, dt + 60t^9\, dt = \left(3t^3 - 4t^7 + 6t^{10}\right)\Big|_0^1 = 5. \tag{4.3.3}$$

□

• Example 4.3.2

Let us redo the previous example with a contour that consists of three "dog legs," namely straight lines from $(0, 0, 0)$ to $(1, 0, 0)$, from $(1, 0, 0)$ to $(1, 1, 0)$, and from $(1, 1, 0)$ to $(1, 1, 1)$. See Figure 4.3.2.

In this particular problem we break the integration down into integrals along each of the legs:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r} + \int_{C_3} \mathbf{F} \cdot d\mathbf{r}. \tag{4.3.4}$$

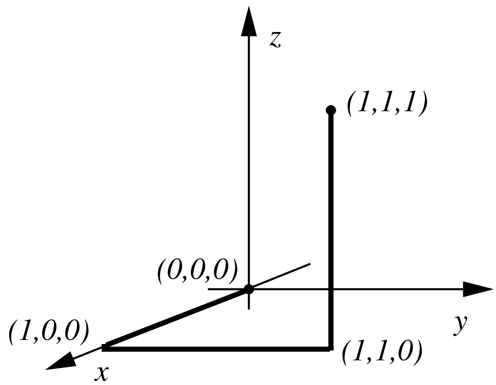


Figure 4.3.2: Diagram for the line integration in Example 4.3.2.

For C_1, $y = z = dy = dz = 0$, and

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (3x^2 + 6 \cdot 0)\, dx - 14 \cdot 0 \cdot 0 \cdot 0 + 20x \cdot 0^2 \cdot 0 = \int_0^1 3x^2\, dx = 1. \qquad (\mathbf{4.3.5})$$

For C_2, $x = 1$ and $z = dx = dz = 0$, so that

$$\int_{C_2} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (3 \cdot 1^2 + 6y) \cdot 0 - 14y \cdot 0 \cdot dy + 20 \cdot 1 \cdot 0^2 \cdot 0 = 0. \qquad (\mathbf{4.3.6})$$

For C_3, $x = y = 1$ and $dx = dy = 0$, so that

$$\int_{C_3} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (3 \cdot 1^2 + 6 \cdot 1) \cdot 0 - 14 \cdot 1 \cdot z \cdot 0 + 20 \cdot 1 \cdot z^2\, dz = \int_0^1 20z^2\, dz = \tfrac{20}{3}. \qquad (\mathbf{4.3.7})$$

Therefore,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \tfrac{23}{3}. \qquad (\mathbf{4.3.8})$$

□

• **Example 4.3.3**

For our third calculation, we redo the first example where the contour is a straight line. The parameterization in this case is $x = y = z = t$ with $0 \le t \le 1$. See Figure 4.3.3. Then,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (3t^2 + 6t)\, dt - 14(t)(t)\, dt + 20t(t)^2\, dt \qquad (\mathbf{4.3.9})$$

$$= \int_0^1 (3t^2 + 6t - 14t^2 + 20t^3)\, dt = \tfrac{13}{3}. \qquad (\mathbf{4.3.10})$$

□

An interesting aspect of these three examples is that, although we used a common vector field and moved from $(0,0,0)$ to $(1,1,1)$ in each case, we obtained a different answer in each case. Thus, for this vector field, the line integral is *path dependent*. This is generally

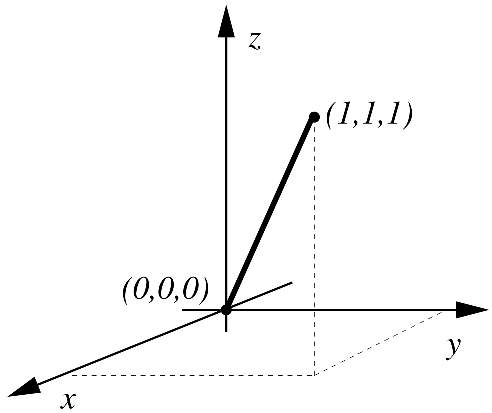


Figure 4.3.3: Diagram for the line integration in Example 4.3.3.

true. In the next section we will meet *conservative vector fields* where the results will be path independent.

• Example 4.3.4

If $\mathbf{F} = (x^2 + y^2)\mathbf{i} - 2xy\mathbf{j} + x\mathbf{k}$, let us evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$ if the contour is that portion of the circle $x^2 + y^2 = a^2$ from the point $(a, 0, 3)$ to $(-a, 0, 3)$. See Figure 4.3.4.

The parametric equations for this example are $x = a\cos(\theta)$, $dx = -a\ \sin(\theta)\,d\theta$, $y = a\sin(\theta)$, $dy = a\cos(\theta)\,d\theta$, $z = 3$, and $dz = 0$ with $0 \le \theta \le \pi$. Therefore,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^\pi [a^2\cos^2(\theta) + a^2\sin^2(\theta)][-a\sin(\theta)\,d\theta] - 2a^2\cos(\theta)\sin(\theta)[a\cos(\theta)\,d\theta] + a\cos(\theta)\cdot 0 \tag{4.3.11}$$

$$= -a^3\int_0^\pi \sin(\theta)\,d\theta - 2a^3\int_0^\pi \cos^2(\theta)\sin(\theta)\,d\theta \tag{4.3.12}$$

$$= a^3\cos(\theta)\Big|_0^\pi + \tfrac{2}{3}a^3\cos^3(\theta)\Big|_0^\pi = -2a^3 - \tfrac{4}{3}a^3 = -\tfrac{10}{3}a^3. \tag{4.3.13}$$

□

• Example 4.3.5: Circulation

Let $\mathbf{v}(x, y, z)$ denote the velocity at the point (x, y, z) in a moving fluid. If it varies with time, this is the velocity at a particular instant of time. The integral $\oint_C \mathbf{v}\cdot d\mathbf{r}$ around a closed path C is called the *circulation* around that path. The average component of velocity along the path is

$$\overline{v}_s = \frac{\oint_C v_s\,ds}{s} = \frac{\oint_C \mathbf{v}\cdot d\mathbf{r}}{s}, \tag{4.3.14}$$

where s is the total length of the path. The circulation is thus $\oint_C \mathbf{v}\cdot d\mathbf{r} = \overline{v}_s s$, the product of the length of the path and the average velocity along the path. When the circulation is positive, the flow is more in the direction of integration than opposite to it. Circulation is thus an indication and to some extent a measure of motion around the path.

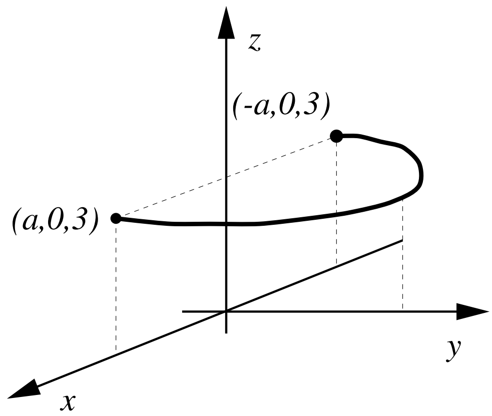


Figure 4.3.4: Diagram for the line integration in Example 4.3.4.

Problems

Evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$ for the following vector fields and curves:

1. $\mathbf{F} = y\sin(\pi z)\mathbf{i} + x^2e^y\mathbf{j} + 3xz\mathbf{k}$ and C is the curve $x = t$, $y = t^2$, and $z = t^3$ from $(0,0,0)$ to $(1,1,1)$. Use MATLAB to illustrate the parametric curves.

2. $\mathbf{F} = y\mathbf{i} + z\mathbf{j} + x\mathbf{k}$ and C consists of the line segments from $(0,0,0)$ to $(2,3,0)$, and from $(2,3,0)$ to $(2,3,4)$. Use MATLAB to illustrate the parametric curves.

3. $\mathbf{F} = e^x\mathbf{i} + xe^{xy}\mathbf{j} + xye^{xyz}\mathbf{k}$ and C is the curve $x = t$, $y = t^2$, and $z = t^3$ with $0 \le t \le 2$. Use MATLAB to illustrate the parametric curves.

4. $\mathbf{F} = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}$ and C is the curve $x = t^3$, $y = t^2$, and $z = t$ with $1 \le t \le 2$. Use MATLAB to illustrate the parametric curves.

5. $\mathbf{F} = y\mathbf{i} - x\mathbf{j} + 3xy\mathbf{k}$ and C consists of the semicircle $x^2 + y^2 = 4$, $z = 0$, $y > 0$, and the line segment from $(-2,0,0)$ to $(2,0,0)$. Use MATLAB to illustrate the parametric curves.

6. $\mathbf{F} = (x + 2y)\mathbf{i} + (6y - 2x)\mathbf{j}$ and C consists of the sides of the triangle with vertices at $(0,0,0)$, $(1,1,1)$, and $(1,1,0)$. Proceed from $(0,0,0)$ to $(1,1,1)$ to $(1,1,0)$ and back to $(0,0,0)$. Use MATLAB to illustrate the parametric curves.

7. $\mathbf{F} = 2xz\mathbf{i} + 4y^2\mathbf{j} + x^2\mathbf{k}$ and C is taken counterclockwise around the ellipse $x^2/4 + y^2/9 = 1$, $z = 1$. Use MATLAB to illustrate the parametric curves.

8. $\mathbf{F} = 2x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and C is the contour $x = t$, $y = \sin(t)$, and $z = \cos(t) + \sin(t)$ with $0 \le t \le 2\pi$. Use MATLAB to illustrate the parametric curves.

9. $\mathbf{F} = (2y^2 + z)\mathbf{i} + 4xy\mathbf{j} + x\mathbf{k}$ and C is the spiral $x = \cos(t)$, $y = \sin(t)$, and $z = t$ with $0 \le t \le 2\pi$ between the points $(1,0,0)$ and $(1,0,2\pi)$. Use MATLAB to illustrate the parametric curves.

10. $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} + (z^2 + 2xy)\mathbf{k}$ and C consists of the edges of the triangle with vertices at $(0,0,0)$, $(1,1,0)$, and $(0,1,0)$. Proceed from $(0,0,0)$ to $(1,1,0)$ to $(0,1,0)$ and back to $(0,0,0)$. Use MATLAB to illustrate the parametric curves.

4.4 THE POTENTIAL FUNCTION

In Section 4.2 we showed that the curl operation applied to a gradient produces the zero vector: $\nabla \times \nabla\varphi = \mathbf{0}$. Consequently, if we have a vector field $\mathbf{F}$ such that $\nabla \times \mathbf{F} \equiv \mathbf{0}$ everywhere, then that vector field is called a *conservative* field and we can compute a potential φ such that $\mathbf{F} = \nabla\varphi$.

• Example 4.4.1

Let us show that the vector field $\mathbf{F} = ye^{xy}\cos(z)\mathbf{i} + xe^{xy}\cos(z)\mathbf{j} - e^{xy}\sin(z)\mathbf{k}$ is conservative and then find the corresponding potential function.

To show that the field is conservative, we compute the curl of $\mathbf{F}$ or

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ ye^{xy}\cos(z) & xe^{xy}\cos(z) & -e^{xy}\sin(z) \end{vmatrix} = \mathbf{0}. \tag{4.4.1}$$

To find the potential we must solve three partial differential equations:

$$\varphi_x = ye^{xy}\cos(z) = \mathbf{F}\cdot\mathbf{i}, \tag{4.4.2}$$

$$\varphi_y = xe^{xy}\cos(z) = \mathbf{F}\cdot\mathbf{j}, \tag{4.4.3}$$

and

$$\varphi_z = -e^{xy}\sin(z) = \mathbf{F}\cdot\mathbf{k}. \tag{4.4.4}$$

We begin by integrating any one of these three equations. Choosing Equation 4.4.2,

$$\varphi(x,y,z) = e^{xy}\cos(z) + f(y,z). \tag{4.4.5}$$

To find $f(y,z)$ we differentiate Equation 4.4.5 with respect to y and find that

$$\varphi_y = xe^{xy}\cos(z) + f_y(y,z) = xe^{xy}\cos(z) \tag{4.4.6}$$

from Equation 4.4.3. Thus, $f_y = 0$ and $f(y,z)$ can only be a function of z, say $g(z)$. Then,

$$\varphi(x,y,z) = e^{xy}\cos(z) + g(z). \tag{4.4.7}$$

Finally,

$$\varphi_z = -e^{xy}\sin(z) + g'(z) = -e^{xy}\sin(z) \tag{4.4.8}$$

from Equation 4.4.4 and $g'(z) = 0$. Therefore, the potential is

$$\varphi(x,y,z) = e^{xy}\cos(z) + \text{constant}. \tag{4.4.9}$$

□

Potentials can be very useful in computing line integrals, because

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C \varphi_x\,dx + \varphi_y\,dy + \varphi_z\,dz = \int_C d\varphi = \varphi(B) - \varphi(A), \tag{4.4.10}$$

where the point B is the terminal point of the integration while the point A is the starting point. Thus, any path integration between any two points is *path independent*.

Finally, if we close the path so that A and B coincide, then

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0. \tag{4.4.11}$$

It should be noted that the converse is *not* true. Just because $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$, we do not necessarily have a conservative field $\mathbf{F}$.

In summary then, an irrotational vector in a given region has three fundamental properties: (1) its integral around every simply connected circuit is zero, (2) its curl equals zero, and (3) it is the gradient of a scalar function. For continuously differentiable vectors, these properties are equivalent. For vectors that are only piece-wise differentiable, this is not true. Generally the first property is the most fundamental and is taken as the definition of irrotationality.

• Example 4.4.2

Using the potential found in Example 4.4.1, let us find the value of the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ from the point $(0, 0, 0)$ to $(-1, 2, \pi)$.

From Equation 4.4.9,

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \left[e^{xy} \cos(z) + \text{constant}\right]\Big|_{(0,0,0)}^{(-1,2,\pi)} = -1 - e^{-2}. \tag{4.4.12}$$

Problems

Verify that the following vector fields are conservative and then find the corresponding potential:

1. $\mathbf{F} = 2xy\mathbf{i} + (x^2 + 2yz)\mathbf{j} + (y^2 + 4)\mathbf{k}$
2. $\mathbf{F} = (2x + 2ze^{2x})\mathbf{i} + (2y - 1)\mathbf{j} + e^{2x}\mathbf{k}$
3. $\mathbf{F} = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}$
4. $\mathbf{F} = 2x\mathbf{i} + 3y^2\mathbf{j} + 4z^3\mathbf{k}$
5. $\mathbf{F} = [2x\sin(y) + e^{3z}]\mathbf{i} + x^2\cos(y)\mathbf{j} + (3xe^{3z} + 4)\mathbf{k}$
6. $\mathbf{F} = (2x + 5)\mathbf{i} + 3y^2\mathbf{j} + (1/z)\mathbf{k}$
7. $\mathbf{F} = e^{2z}\mathbf{i} + 3y^2\mathbf{j} + 2xe^{2z}\mathbf{k}$
8. $\mathbf{F} = y\mathbf{i} + (x + z)\mathbf{j} + y\mathbf{k}$
9. $\mathbf{F} = (z + y)\mathbf{i} + x\mathbf{j} + x\mathbf{k}$.

4.5 SURFACE INTEGRALS

Surface integrals appear in such diverse fields as electromagnetism and fluid mechanics. For example, if we were oceanographers we might be interested in the rate of volume of seawater through an instrument that has the curved surface S. The volume rate equals $\iint_S \mathbf{v} \cdot \mathbf{n}\, d\sigma$, where $\mathbf{v}$ is the velocity and $\mathbf{n}\, d\sigma$ is an infinitesimally small element on the surface of the instrument. The surface element $\mathbf{n}\, d\sigma$ must have an orientation (given by $\mathbf{n}$) because it makes a considerable difference whether the flow is directly through the surface

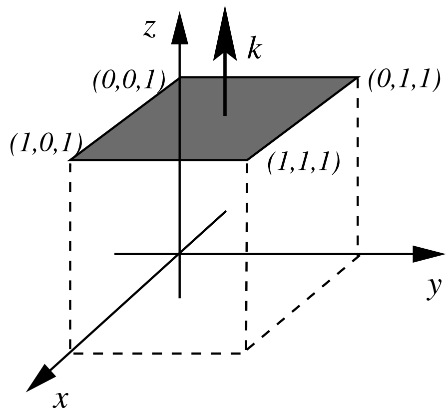


Figure 4.5.1: Diagram for the surface integration in Example 4.5.1.

or at right angles. In the special case when the surface encloses a three-dimensional volume, then we have a *closed surface integral*.

To illustrate the concept of computing a surface integral, we will do three examples with simple geometries. Later we will show how to use surface coordinates to do more complicated geometries.

• **Example 4.5.1**

Let us find the flux out the top of a unit cube if the vector field is $\mathbf{F} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. See Figure 4.5.1.

The top of a unit cube consists of the surface $z = 1$ with $0 \leq x \leq 1$ and $0 \leq y \leq 1$. By inspection the unit normal to this surface is $\mathbf{n} = \mathbf{k}$, or $\mathbf{n} = -\mathbf{k}$. Because we are interested in the flux *out* of the unit cube, $\mathbf{n} = \mathbf{k}$, and

$$\iint_S \mathbf{F} \cdot \mathbf{n}\, d\sigma = \int_0^1 \int_0^1 (x\mathbf{i} + y\mathbf{j} + \mathbf{k}) \cdot \mathbf{k}\, dx\, dy = 1, \qquad \textbf{(4.5.1)}$$

because $z = 1$. □

• **Example 4.5.2**

Let us find the flux out of that portion of the cylinder $y^2 + z^2 = 4$ in the first octant bounded by $x = 0$, $x = 3$, $y = 0$, and $z = 0$. The vector field is $\mathbf{F} = x\mathbf{i} + 2z\mathbf{j} + y\mathbf{k}$. See Figure 4.5.2.

Because we are dealing with a cylinder, cylindrical coordinates are appropriate. Let $y = 2\cos(\theta)$, $z = 2\sin(\theta)$, and $x = x$ with $0 \leq \theta \leq \pi/2$. To find $\mathbf{n}$, we use the gradient in conjunction with the definition of the surface of the cylinder $f(x, y, z) = y^2 + z^2 = 4$. Then,

$$\mathbf{n} = \frac{\nabla f}{|\nabla f|} = \frac{2y\mathbf{j} + 2z\mathbf{k}}{\sqrt{4y^2 + 4z^2}} = \frac{y}{2}\mathbf{j} + \frac{z}{2}\mathbf{k}, \qquad \textbf{(4.5.2)}$$

because $y^2 + z^2 = 4$ along the surface. Since we want the flux *out* of the surface, then $\mathbf{n} = y\mathbf{j}/2 + z\mathbf{k}/2$, whereas the flux *into* the surface would require $\mathbf{n} = -y\mathbf{j}/2 - z\mathbf{k}/2$. Therefore,

$$\mathbf{F} \cdot \mathbf{n} = (x\mathbf{i} + 2z\mathbf{j} + y\mathbf{k}) \cdot \left(\frac{y}{2}\mathbf{j} + \frac{z}{2}\mathbf{k}\right) = \frac{3yz}{2} = 6\cos(\theta)\sin(\theta). \qquad \textbf{(4.5.3)}$$

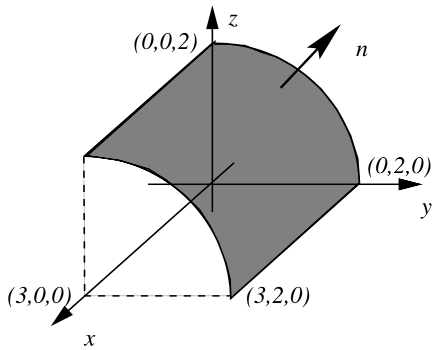


Figure 4.5.2: Diagram for the surface integration in Example 4.5.2.

What is $d\sigma$? Our infinitesimal surface area has a side in the x direction of length dx and a side in the θ direction of length $2\,d\theta$ because the radius equals 2. Therefore, $d\sigma = 2\,dx\,d\theta$.

Bringing all of these elements together,

$$\iint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^3 \int_0^{\pi/2} 12\cos(\theta)\sin(\theta)\,d\theta\,dx = 6\int_0^3 \left[\sin^2(\theta)\Big|_0^{\pi/2}\right]\,dx = 6\int_0^3 dx = 18. \tag{4.5.4}$$

As counterpoint to this example, let us find the flux out of the pie-shaped surface at $x = 3$. In this case, $y = r\cos(\theta)$, $z = r\sin(\theta)$, and

$$\iint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^{\pi/2}\int_0^2 [3\mathbf{i} + 2r\sin(\theta)\mathbf{j} + r\cos(\theta)\mathbf{k}]\cdot\mathbf{i}\,r\,dr\,d\theta = 3\int_0^{\pi/2}\int_0^2 r\,dr\,d\theta = 3\pi. \tag{4.5.5}$$

□

• Example 4.5.3

Let us find the flux of the vector field $\mathbf{F} = y^2\mathbf{i} + x^2\mathbf{j} + 5z\mathbf{k}$ out of the hemispheric surface $x^2 + y^2 + z^2 = a^2$, $z > 0$. See Figure 4.5.3.

We begin by finding the outwardly pointing normal. Because the surface is defined by $f(x, y, z) = x^2 + y^2 + z^2 = a^2$,

$$\mathbf{n} = \frac{\nabla f}{|\nabla f|} = \frac{2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}}{\sqrt{4x^2 + 4y^2 + 4z^2}} = \frac{x}{a}\mathbf{i} + \frac{y}{a}\mathbf{j} + \frac{z}{a}\mathbf{k}, \tag{4.5.6}$$

because $x^2 + y^2 + z^2 = a^2$. This is also the outwardly pointing normal since $\mathbf{n} = \mathbf{r}/a$, where $\mathbf{r}$ is the radial vector.

Using spherical coordinates, $x = a\cos(\varphi)\sin(\theta)$, $y = a\sin(\varphi)\sin(\theta)$, and $z = a\cos(\theta)$, where φ is the angle made by the projection of the point onto the equatorial plane, measured from the x-axis, and θ is the colatitude or "cone angle" measured from the z-axis. To compute $d\sigma$, the infinitesimal length in the θ direction is $a\,d\theta$ while in the φ direction it is $a\sin(\theta)\,d\varphi$, where the $\sin(\theta)$ factor takes into account the convergence of the meridians. Therefore, $d\sigma = a^2\sin(\theta)\,d\theta\,d\varphi$, and

$$\iint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^{2\pi}\int_0^{\pi/2} \left(y^2\mathbf{i} + x^2\mathbf{j} + 5z\mathbf{k}\right)\left(\frac{x}{a}\mathbf{i} + \frac{y}{a}\mathbf{j} + \frac{z}{a}\mathbf{k}\right) a^2\sin(\theta)\,d\theta\,d\varphi \tag{4.5.7}$$

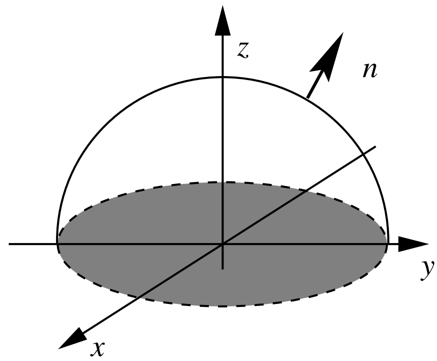


Figure 4.5.3: Diagram for the surface integration in Example 4.5.3.

$$\iint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^{2\pi}\int_0^{\pi/2}\left(\frac{xy^2}{a}+\frac{x^2y}{a}+\frac{5z^2}{a}\right)a^2\sin(\theta)\,d\theta\,d\varphi \qquad (4.5.8)$$

$$= \int_0^{\pi/2}\int_0^{2\pi}\big[a^4\cos(\varphi)\sin^2(\varphi)\sin^4(\theta) + a^4\cos^2(\varphi)\sin(\varphi)\sin^4(\theta) + 5a^3\cos^2(\theta)\sin(\theta)\big]\,d\varphi\,d\theta \qquad (4.5.9)$$

$$= \int_0^{\pi/2}\left[\frac{a^4}{3}\sin^3(\varphi)\Big|_0^{2\pi}\sin^4(\theta) - \frac{a^4}{3}\cos^3(\varphi)\Big|_0^{2\pi}\sin^4(\theta) + 5a^3\cos^2(\theta)\sin(\theta)\varphi\Big|_0^{2\pi}\right]d\theta \qquad (4.5.10)$$

$$= 10\pi a^3\int_0^{\pi/2}\cos^2(\theta)\sin(\theta)\,d\theta = -\frac{10\pi a^3}{3}\cos^3(\theta)\Big|_0^{\pi/2} = \frac{10\pi a^3}{3}. \qquad (4.5.11)$$

□

Although these techniques apply for simple geometries such as a cylinder or sphere, we would like a *general* method for treating any arbitrary surface. We begin by noting that a surface is an aggregate of points whose coordinates are functions of two variables. For example, in the previous example, the surface was described by the coordinates φ and θ. Let us denote these surface coordinates in general by u and v. Consequently, on any surface we can reexpress x, y, and z in terms of u and v: $x = x(u,v)$, $y = y(u,v)$, and $z = z(u,v)$.

Next, we must find an infinitesimal element of area. The position vector to the surface is $\mathbf{r} = x(u,v)\mathbf{i} + y(u,v)\mathbf{j} + z(u,v)\mathbf{k}$. Therefore, the tangent vectors along $v =$ constant, $\mathbf{r}_u$, and along $u =$ constant, $\mathbf{r}_v$, equal

$$\mathbf{r}_u = x_u\mathbf{i} + y_u\mathbf{j} + z_u\mathbf{k}, \qquad (4.5.12)$$

and

$$\mathbf{r}_v = x_v\mathbf{i} + y_v\mathbf{j} + z_v\mathbf{k}. \qquad (4.5.13)$$

Consequently, the sides of the infinitesimal area are $\mathbf{r}_u\,du$ and $\mathbf{r}_v\,dv$. Therefore, the vectorial area of the parallelogram that these vectors form is

$$\mathbf{n}\,d\sigma = \mathbf{r}_u \times \mathbf{r}_v\,du\,dv \qquad (4.5.14)$$

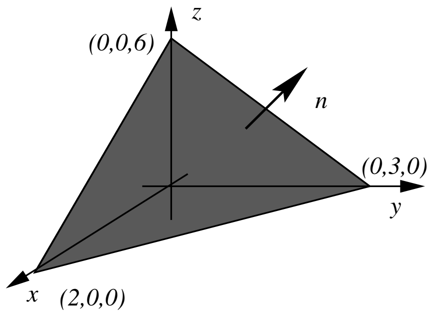


Figure 4.5.4: Diagram for the surface integration in Example 4.5.4.

and is called the *vector element of area* on the surface. Thus, we may convert $\mathbf{F}\cdot\mathbf{n}\,d\sigma$ into an expression involving only u and v and then evaluate the surface integral by integrating over the appropriate domain in the uv-plane. Of course, we are in trouble if $\mathbf{r}_u\times\mathbf{r}_v=\mathbf{0}$. Therefore, we only treat regular points where $\mathbf{r}_u\times\mathbf{r}_v\neq\mathbf{0}$. In the next few examples, we show how to use these surface coordinates to evaluate surface integrals.

• Example 4.5.4

Let us find the flux of the vector field $\mathbf{F}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ through the top of the plane $3x+2y+z=6$, which lies in the first octant. See Figure 4.5.4.

Our parametric equations are $x=u$, $y=v$, and $z=6-3u-2v$. Therefore,

$$\mathbf{r}=u\mathbf{i}+v\mathbf{j}+(6-3u-2v)\mathbf{k}, \tag{4.5.15}$$

so that

$$\mathbf{r}_u=\mathbf{i}-3\mathbf{k},\quad \mathbf{r}_v=\mathbf{j}-2\mathbf{k}, \tag{4.5.16}$$

and

$$\mathbf{r}_u\times\mathbf{r}_v=3\mathbf{i}+2\mathbf{j}+\mathbf{k}. \tag{4.5.17}$$

Bring all of these elements together,

$$\iint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma=\int_0^2\int_0^{3-3u/2}(3u+2v+6-3u-2v)\,dv\,du=6\int_0^2\int_0^{3-3u/2}dv\,du \tag{4.5.18}$$

$$=6\int_0^2(3-3u/2)\,du=6\left(3u-\tfrac{3}{4}u^2\right)\Big|_0^2=18. \tag{4.5.19}$$

To set up the limits of integration, we note that the area in u,v space corresponds to the xy-plane. On the xy-plane, $z=0$ and $3u+2v=6$, along with boundaries $u=v=0$. □

• Example 4.5.5

Let us find the flux of the vector field $\mathbf{F}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ through the top of the surface $z=xy+1$, which covers the square $0\le x\le 1$, $0\le y\le 1$ in the xy-plane. See Figure 4.5.5.

Our parametric equations are $x=u$, $y=v$, and $z=uv+1$ with $0\le u\le 1$ and $0\le v\le 1$. Therefore,

$$\mathbf{r}=u\mathbf{i}+v\mathbf{j}+(uv+1)\mathbf{k}, \tag{4.5.20}$$

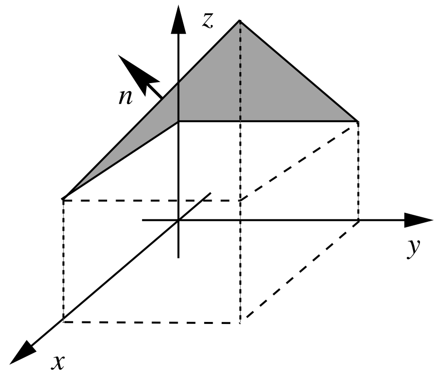


Figure 4.5.5: Diagram for the surface integration in Example 4.5.5.

so that

$$\mathbf{r}_u = \mathbf{i} + v\mathbf{k}, \qquad \mathbf{r}_v = \mathbf{j} + u\mathbf{k}, \tag{4.5.21}$$

and

$$\mathbf{r}_u \times \mathbf{r}_v = -v\mathbf{i} - u\mathbf{j} + \mathbf{k}. \tag{4.5.22}$$

Bring all of these elements together,

$$\iint_S \mathbf{F} \cdot \mathbf{n}\, d\sigma = \int_0^1 \int_0^1 [u\mathbf{i} + v\mathbf{j} + (uv+1)\mathbf{k}] \cdot (-v\mathbf{i} - u\mathbf{j} + \mathbf{k})\, du\, dv \tag{4.5.23}$$

$$= \int_0^1 \int_0^1 (1 - uv)\, du\, dv = \int_0^1 \left(u - \tfrac{1}{2}u^2 v\right)\Big|_0^1 dv \tag{4.5.24}$$

$$= \int_0^1 \left(1 - \tfrac{1}{2}v\right)\, dv = \left(v - \tfrac{1}{4}v^2\right)\Big|_0^1 = \tfrac{3}{4}. \tag{4.5.25}$$

□

• Example 4.5.6

Let us find the flux of the vector field $\mathbf{F} = 4xz\mathbf{i} + xyz^2\mathbf{j} + 3z\mathbf{k}$ through the exterior surface of the cone $z^2 = x^2 + y^2$ above the xy-plane and below $z = 4$. See Figure 4.5.6.

A natural choice for the surface coordinates is polar coordinates r and θ. Because $x = r\cos(\theta)$ and $y = r\sin(\theta)$, $z = r$. Then,

$$\mathbf{r} = r\cos(\theta)\mathbf{i} + r\sin(\theta)\mathbf{j} + r\mathbf{k} \tag{4.5.26}$$

with $0 \le r \le 4$ and $0 \le \theta \le 2\pi$ so that

$$\mathbf{r}_r = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j} + \mathbf{k}\mathbf{r}_\theta = -r\sin(\theta)\mathbf{i} + r\cos(\theta)\mathbf{j}, \tag{4.5.27}$$

and

$$\mathbf{r}_r \times \mathbf{r}_\theta = -r\cos(\theta)\mathbf{i} - r\sin(\theta)\mathbf{j} + r\mathbf{k}. \tag{4.5.28}$$

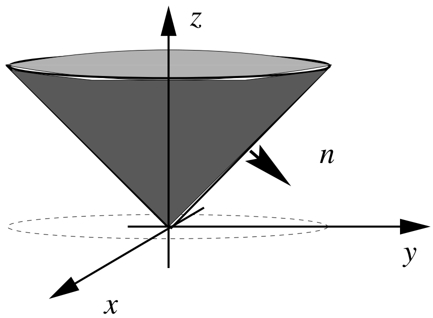


Figure 4.5.6: Diagram for the surface integration in Example 4.5.6.

This is the unit area *inside* the cone. Because we want the exterior surface, we must take the negative of Equation 4.5.28. Bring all of these elements together,

$$\iint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^4 \int_0^{2\pi} \{[4r\cos(\theta)]r[r\cos(\theta)] + [r^2\sin(\theta)\cos(\theta)]r^2[r\sin(\theta)] - 3r^2\}\,d\theta\,dr \tag{4.5.29}$$

$$= \int_0^4 \left\{ 2r^3\left[\theta + \tfrac{1}{2}\sin(2\theta)\right]\Big|_0^{2\pi} + r^5\tfrac{1}{3}\sin^3(\theta)\Big|_0^{2\pi} - 3r^2\theta\Big|_0^{2\pi} \right\} dr \tag{4.5.30}$$

$$= \int_0^4 \left(4\pi r^3 - 6\pi r^2\right)\,dr = \left(\pi r^4 - 2\pi r^3\right)\Big|_0^4 = 128\pi. \tag{4.5.31}$$

Problems

Compute the surface integral $\iint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma$ for the following vector fields and surfaces:

1. $\mathbf{F} = x\mathbf{i} - z\mathbf{j} + y\mathbf{k}$ and the surface is the top side of the $z = 1$ plane where $0 \le x \le 1$ and $0 \le y \le 1$.

2. $\mathbf{F} = x\mathbf{i} + y\mathbf{j} + xz\mathbf{k}$ and the surface is the top side of the cylinder $x^2 + y^2 = 9$, $z = 0$, and $z = 1$.

3. $\mathbf{F} = xy\mathbf{i} + z\mathbf{j} + xz\mathbf{k}$ and the surface consists of both exterior *ends* of the cylinder defined by $x^2 + y^2 = 4$, $z = 0$, and $z = 2$.

4. $\mathbf{F} = x\mathbf{i} + z\mathbf{j} + y\mathbf{k}$ and the surface is the lateral and exterior sides of the cylinder defined by $x^2 + y^2 = 4$, $z = -3$, and $z = 3$.

5. $\mathbf{F} = xy\mathbf{i} + z^2\mathbf{j} + y\mathbf{k}$ and the surface is the curved exterior side of the cylinder $y^2 + z^2 = 9$ in the first octant bounded by $x = 0$, $x = 1$, $y = 0$, and $z = 0$.

6. $\mathbf{F} = y\mathbf{j} + z^2\mathbf{k}$ and the surface is the exterior of the semicircular cylinder $y^2 + z^2 = 4$, $z \ge 0$, cut by the planes $x = 0$ and $x = 1$.

7. $\mathbf{F} = z\mathbf{i} + x\mathbf{j} + y\mathbf{k}$ and the surface is the curved exterior side of the cylinder $x^2 + y^2 = 4$ in the first octant cut by the planes $z = 1$ and $z = 2$.

8. $\mathbf{F} = x^2\mathbf{i} - z^2\mathbf{j} + yz\mathbf{k}$ and the surface is the exterior of the hemispheric surface of $x^2 + y^2 + z^2 = 16$ above the plane $z = 2$.

9. $\mathbf{F} = y\mathbf{i} + x\mathbf{j} + y\mathbf{k}$ and the surface is the top of the surface $z = x + 1$, where $-1 \le x \le 1$ and $-1 \le y \le 1$.

10. $\mathbf{F} = z\mathbf{i} + x\mathbf{j} - 3z\mathbf{k}$ and the surface is the top side of the plane $x + y + z = 2a$ that lies above the square $0 \le x \le a$, $0 \le y \le a$ in the xy-plane.

11. $\mathbf{F} = (y^2 + z^2)\mathbf{i} + (x^2 + z^2)\mathbf{j} + (x^2 + y^2)\mathbf{k}$ and the surface is the top side of the surface $z = 1 - x^2$ with $-1 \le x \le 1$ and $-2 \le y \le 2$.

12. $\mathbf{F} = y^2\mathbf{i} + xz\mathbf{j} - \mathbf{k}$ and the surface is the cone $z = \sqrt{x^2 + y^2}$, $0 \le z \le 1$, with the normal pointing away from the z-axis.

13. $\mathbf{F} = y^2\mathbf{i} + x^2\mathbf{j} + 5z\mathbf{k}$ and the surface is the top side of the plane $z = y + 1$, where $-1 \le x \le 1$ and $-1 \le y \le 1$.

14. $\mathbf{F} = -y\mathbf{i} + x\mathbf{j} + z\mathbf{k}$ and the surface is the exterior or bottom side of the paraboloid $z = x^2 + y^2$, where $0 \le z \le 1$.

15. $\mathbf{F} = -y\mathbf{i} + x\mathbf{j} + 6z^2\mathbf{k}$ and the surface is the exterior of the paraboloids $z = 4 - x^2 - y^2$ and $z = x^2 + y^2$.

4.6 GREEN'S LEMMA

Consider a rectangle in the xy-plane that is bounded by the lines $x = a$, $x = b$, $y = c$, and $y = d$. We assume that the boundary of the rectangle is a piece-wise smooth curve that we denote by C. If we have a continuously differentiable vector function $\mathbf{F} = P(x,y)\mathbf{i} + Q(x,y)\mathbf{j}$ at each point of enclosed region R, then

$$\iint_R \frac{\partial Q}{\partial x}\,dA = \int_c^d \left[\int_a^b \frac{\partial Q}{\partial x}\,dx\right] dy = \int_c^d Q(b,y)\,dy - \int_c^d Q(a,y)\,dy = \oint_C Q(x,y)\,dy, \tag{4.6.1}$$

where the last integral is a closed line integral counterclockwise around the rectangle because the horizontal sides vanish, since $dy = 0$. By similar arguments,

$$\iint_R \frac{\partial P}{\partial y}\,dA = -\oint_C P(x,y)\,dx \tag{4.6.2}$$

so that

$$\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA = \oint_C P(x,y)\,dx + Q(x,y)\,dy. \tag{4.6.3}$$

This result, often known as *Green's lemma*, may be expressed in vector form as

$$\boxed{\oint_C \mathbf{F}\cdot d\mathbf{r} = \iint_R \nabla\times\mathbf{F}\cdot\mathbf{k}\,dA.} \tag{4.6.4}$$

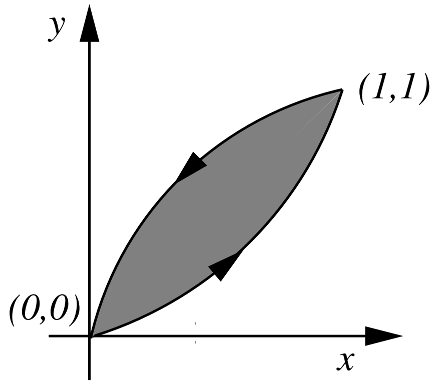


Figure 4.6.1: Diagram for the verification of Green's lemma in Example 4.6.1.

Although this proof was for a rectangular area, it can be generalized to *any* simply closed region on the xy-plane as follows. Consider an area that is surrounded by simply closed curves. Within the closed contour we can divide the area into an infinite number of infinitesimally small rectangles and apply Equation 4.6.4 to each rectangle. When we sum up all of these rectangles, we find $\iint_R \nabla \times \mathbf{F} \cdot \mathbf{k}\, dA$, where the integration is over the entire surface area. On the other hand, away from the boundary, the line integral along any one edge of a rectangle cancels the line integral along the same edge in a contiguous rectangle. Thus, the only nonvanishing contribution from the line integrals arises from the outside boundary of the domain $\oint_C \mathbf{F} \cdot d\mathbf{r}$.

• Example 4.6.1

Let us *verify* Green's lemma using the vector field $\mathbf{F} = (3x^2 - 8y^2)\mathbf{i} + (4y - 6xy)\mathbf{j}$, and the enclosed area lies between the curves $y = \sqrt{x}$ and $y = x^2$. The two curves intersect at $x = 0$ and $x = 1$. See Figure 4.6.1.

We begin with the line integral:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (3x^2 - 8x^4)\, dx + (4x^2 - 6x^3)(2x\, dx)$$

$$+ \int_1^0 (3x^2 - 8x)\, dx + (4x^{1/2} - 6x^{3/2})(\tfrac{1}{2}x^{-1/2}\, dx) \tag{4.6.5}$$

$$= \int_0^1 (-20x^4 + 8x^3 + 11x - 2)\, dx = \tfrac{3}{2}. \tag{4.6.6}$$

In Equation 4.6.6 we used $y = x^2$ in the first integral and $y = \sqrt{x}$ in our return integration. For the areal integration,

$$\iint_R \nabla \times \mathbf{F} \cdot \mathbf{k}\, dA = \int_0^1 \int_{x^2}^{\sqrt{x}} 10y\, dy\, dx = \int_0^1 5y^2\Big|_{x^2}^{\sqrt{x}}\, dx = 5\int_0^1 (x - x^4)\, dx = \tfrac{3}{2} \tag{4.6.7}$$

and Green's lemma is verified in this particular case. □

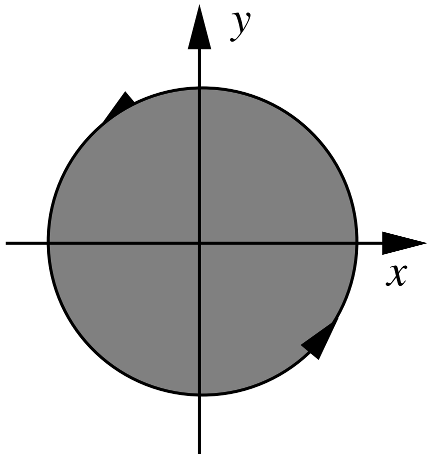


Figure 4.6.2: Diagram for the verification of Green's lemma in Example 4.6.3.

• Example 4.6.2

Let us redo Example 4.6.1 except that the closed contour is the triangular region defined by the lines $x = 0$, $y = 0$, and $x + y = 1$.

The line integral is

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_0^1 (3x^2 - 8 \cdot 0^2)dx + (4 \cdot 0 - 6x \cdot 0) \cdot 0$$

$$+ \int_0^1 [3(1-y)^2 - 8y^2](-dy) + [4y - 6(1-y)y]\, dy$$

$$+ \int_1^0 (3 \cdot 0^2 - 8y^2) \cdot 0 + (4y - 6 \cdot 0 \cdot y)\, dy \quad (\mathbf{4.6.8})$$

$$= \int_0^1 3x^2\, dx - \int_0^1 4y\, dy + \int_0^1 (-3 + 4y + 11y^2)\, dy \quad (\mathbf{4.6.9})$$

$$= x^3\Big|_0^1 - 2y^2\Big|_0^1 + \left(-3y + 2y^2 + \tfrac{11}{3}y^3\right)\Big|_0^1 = \tfrac{5}{3}. \quad (\mathbf{4.6.10})$$

On the other hand, the areal integration is

$$\iint_R \nabla \times \mathbf{F} \cdot \mathbf{k}\, dA = \int_0^1 \int_0^{1-x} 10y\, dy\, dx = \int_0^1 5y^2\Big|_0^{1-x}\, dx \quad (\mathbf{4.6.11})$$

$$= 5 \int_0^1 (1-x)^2\, dx = -\, \tfrac{5}{3}(1-x)^3\Big|_0^1 = \tfrac{5}{3} \quad (\mathbf{4.6.12})$$

and Green's lemma is verified in this particular case. □

• Example 4.6.3

Let us verify Green's lemma using the vector field $\mathbf{F} = (3x + 4y)\mathbf{i} + (2x - 3y)\mathbf{j}$, and the closed contour is a circle of radius two centered at the origin of the xy-plane. See Figure 4.6.2.

Beginning with the line integration,

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi} [6\cos(\theta)+8\sin(\theta)][-2\sin(\theta)\,d\theta] + [4\cos(\theta)-6\sin(\theta)][2\cos(\theta)\,d\theta] \tag{4.6.13}$$

$$= \int_0^{2\pi} [-24\cos(\theta)\sin(\theta) - 16\sin^2(\theta) + 8\cos^2(\theta)]\,d\theta \tag{4.6.14}$$

$$= 12\cos^2(\theta)\Big|_0^{2\pi} - 8\left[\theta - \tfrac{1}{2}\sin(2\theta)\right]\Big|_0^{2\pi} + 4\left[\theta + \tfrac{1}{2}\sin(2\theta)\right]\Big|_0^{2\pi} = -8\pi. \tag{4.6.15}$$

For the areal integration,

$$\iint_R \nabla\times\mathbf{F}\cdot\mathbf{k}\,dA = \int_0^2\int_0^{2\pi} -2\,r\,d\theta\,dr = -8\pi \tag{4.6.16}$$

and Green's lemma is verified in the special case.

Problems

Verify Green's lemma for the following two-dimensional vector fields and contours:

1. $\mathbf{F} = (x^2+4y)\mathbf{i} + (y-x)\mathbf{j}$ and the contour is the square bounded by the lines $x=0$, $y=0$, $x=1$, and $y=1$.

2. $\mathbf{F} = (x-y)\mathbf{i} + xy\mathbf{j}$ and the contour is the square bounded by the lines $x=0$, $y=0$, $x=1$, and $y=1$.

3. $\mathbf{F} = -y^2\mathbf{i} + x^2\mathbf{j}$ and the contour is the triangle bounded by the lines $x=1$, $y=0$, and $y=x$.

4. $\mathbf{F} = (xy-x^2)\mathbf{i} + x^2y\mathbf{j}$ and the contour is the triangle bounded by the lines $y=0$, $x=1$, and $y=x$.

5. $\mathbf{F} = \sin(y)\mathbf{i} + x\cos(y)\mathbf{j}$ and the contour is the triangle bounded by the lines $x+y=1$, $y-x=1$, and $y=0$.

6. $\mathbf{F} = y^2\mathbf{i} + x^2\mathbf{j}$ and the contour is the same contour used in Problem 4.

7. $\mathbf{F} = -y^2\mathbf{i} + x^2\mathbf{j}$ and the contour is the circle $x^2+y^2=4$.

8. $\mathbf{F} = -x^2\mathbf{i} + xy^2\mathbf{j}$ and the contour is the closed circle of radius a.

9. $\mathbf{F} = (6y+x)\mathbf{i} + (y+2x)\mathbf{j}$ and the contour is the circle $(x-1)^2+(y-2)^2=4$.

10. $\mathbf{F} = (x+y)\mathbf{i} + (2x^2-y^2)\mathbf{j}$ and the contour is the boundary of the region determined by the curves $y=x^2$ and $y=4$.

11. $\mathbf{F} = 3y\mathbf{i} + 2x\mathbf{j}$ and the contour is the boundary of the region determined by the curves $y=0$ and $y=\sin(x)$ with $0\le x\le\pi$.

12. $\mathbf{F} = -16y\mathbf{i} + (4e^y + 3x^2)\mathbf{j}$ and the contour is the pie wedge defined by the lines $y = x$, $y = -x$, $x^2 + y^2 = 4$, and $y > 0$.

4.7 STOKES' THEOREM[4]

In Section 4.2 we introduced the vector quantity $\nabla \times \mathbf{v}$, which gives a measure of the rotation of a parcel of fluid lying within the velocity field $\mathbf{v}$. In this section we show how the curl can be used to simplify the calculation of certain closed line integrals.

This relationship between a closed line integral and a surface integral involving the curl is

Stokes' Theorem: *The circulation of* $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ *around the closed boundary* C *of an oriented surface* S *in the direction counterclockwise with respect to the surface's unit normal vector* $\mathbf{n}$ *equals the integral of* $\nabla \times \mathbf{F} \cdot \mathbf{n}$ *over* S*, or*

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_S \nabla \times \mathbf{F} \cdot \mathbf{n}\, d\sigma. \tag{4.7.1}$$

Stokes' theorem requires that all of the functions and derivatives be continuous.

The proof of Stokes' theorem is as follows: Consider a finite surface S whose boundary is the loop C. We divide this surface into a number of small elements $\mathbf{n}\, d\sigma$ and compute the *circulation* $d\Gamma = \oint_L \mathbf{F} \cdot d\mathbf{r}$ around each element. When we add all of the circulations together, the contribution from an integration along a boundary line between two adjoining elements cancels out because the boundary is transversed once in each direction. For this reason, the only contributions that survive are those parts where the element boundaries form part of C. Thus, the sum of all circulations equals $\oint_C \mathbf{F} \cdot d\mathbf{r}$, the circulation around the edge of the whole surface.

Next, let us compute the circulation another way. We begin by finding the Taylor expansion for $P(x, y, z)$ about the arbitrary point (x_0, y_0, z_0):

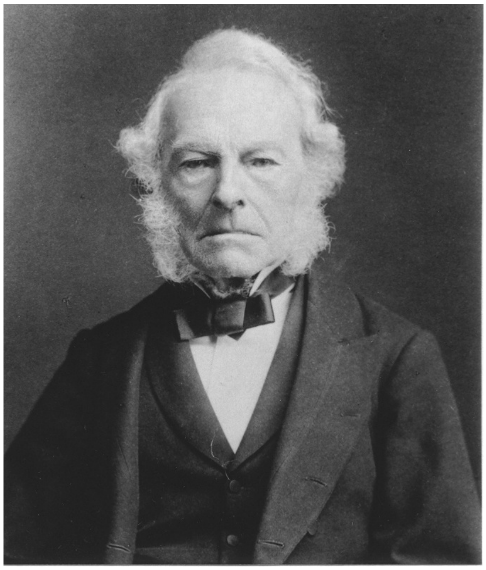
$$\begin{aligned} P(x, y, z) = P(x_0, y_0, z_0) &+ (x - x_0)\frac{\partial P(x_0, y_0, z_0)}{\partial x} + (y - y_0)\frac{\partial P(x_0, y_0, z_0)}{\partial y} \\ &+ (z - z_0)\frac{\partial P(x_0, y_0, z_0)}{\partial z} + \cdots \end{aligned} \tag{4.7.2}$$

with similar expansions for $Q(x, y, z)$ and $R(x, y, z)$. Then

$$\begin{aligned} d\Gamma = \oint_L \mathbf{F} \cdot d\mathbf{r} = P(x_0, y_0, z_0) \oint_L dx &+ \frac{\partial P(x_0, y_0, z_0)}{\partial x} \oint_L (x - x_0)\, dx \\ &+ \frac{\partial P(x_0, y_0, z_0)}{\partial y} \oint_L (y - y_0)\, dy + \cdots + \frac{\partial Q(x_0, y_0, z_0)}{\partial x} \oint_L (x - x_0)\, dy + \cdots, \end{aligned} \tag{4.7.3}$$

where L denotes some small loop located in the surface S. Note that integrals such as $\oint_L dx$ and $\oint_L (x - x_0)\, dx$ vanish.

[4] For the history behind the development of Stokes' theorem, see Katz, V. J., 1979: The history of Stokes' theorem. *Math. Mag.*, **52**, 146–156.



Sir George Gabriel Stokes (1819–1903) was Lucasian Professor of Mathematics at Cambridge University from 1849 until his death. Having learned of an integral theorem from his friend Lord Kelvin, Stokes included it a few years later among his questions on an examination that he wrote for the Smith Prize. It is this integral theorem that we now call Stokes' theorem. (Portrait courtesy of the Royal Society of London.)

If we now require that the loop integrals be in the *clockwise* or *positive* sense so that we preserve the right-hand screw convention, then

$$\mathbf{n}\cdot\mathbf{k}\,\delta\sigma = \oint_L (x-x_0)\,dy = -\oint_L (y-y_0)\,dx, \tag{4.7.4}$$

$$\mathbf{n}\cdot\mathbf{j}\,\delta\sigma = \oint_L (z-z_0)\,dx = -\oint_L (x-x_0)\,dz, \tag{4.7.5}$$

$$\mathbf{n}\cdot\mathbf{i}\,\delta\sigma = \oint_L (y-y_0)\,dz = -\oint_L (z-z_0)\,dy, \tag{4.7.6}$$

and

$$d\Gamma = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i}\cdot\mathbf{n}\,\delta\sigma + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j}\cdot\mathbf{n}\,\delta\sigma + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}\cdot\mathbf{n}\,\delta\sigma = \nabla\times\mathbf{F}\cdot\mathbf{n}\,\delta\sigma. \tag{4.7.7}$$

Therefore, the sum of all circulations in the limit when all elements are made infinitesimally small becomes the surface integral $\iint_S \nabla\times\mathbf{F}\cdot\mathbf{n}\,d\sigma$ and Stokes' theorem is proven. □

In the following examples we first apply Stokes' theorem to a few simple geometries. We then show how to apply this theorem to more complicated surfaces.[5]

[5] Thus, different Stokes for different folks.

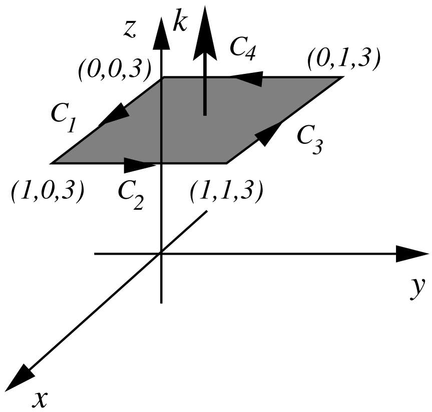


Figure 4.7.1: Diagram for the verification of Stokes' theorem in Example 4.7.1.

• Example 4.7.1

Let us verify Stokes' theorem using the vector field $\mathbf{F} = x^2\mathbf{i} + 2x\mathbf{j} + z^2\mathbf{k}$, and the closed curve is a square with vertices at $(0, 0, 3)$, $(1, 0, 3)$, $(1, 1, 3)$, and $(0, 1, 3)$. See Figure 4.7.1.

We begin with the line integral:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r} + \int_{C_3} \mathbf{F} \cdot d\mathbf{r} + \int_{C_4} \mathbf{F} \cdot d\mathbf{r}, \tag{4.7.8}$$

where C_1, C_2, C_3, and C_4 represent the four sides of the square. Along C_1, x varies while $y = 0$ and $z = 3$. Therefore,

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 x^2\, dx + 2x \cdot 0 + 9 \cdot 0 = \tfrac{1}{3}, \tag{4.7.9}$$

because $dy = dz = 0$, and $z = 3$. Along C_2, y varies with $x = 1$ and $z = 3$. Therefore,

$$\int_{C_2} \mathbf{F} \cdot d\mathbf{r} = \int_0^1 1^2 \cdot 0 + 2 \cdot 1 \cdot dy + 9 \cdot 0 = 2. \tag{4.7.10}$$

Along C_3, x again varies with $y = 1$ and $z = 3$, and so,

$$\int_{C_3} \mathbf{F} \cdot d\mathbf{r} = \int_1^0 x^2\, dx + 2x \cdot 0 + 9 \cdot 0 = -\tfrac{1}{3}. \tag{4.7.11}$$

Note how the limits run from 1 to 0 because x is decreasing. Finally, for C_4, y again varies with $x = 0$ and $z = 3$. Hence,

$$\int_{C_4} \mathbf{F} \cdot d\mathbf{r} = \int_1^0 0^2 \cdot 0 + 2 \cdot 0 \cdot dy + 9 \cdot 0 = 0. \tag{4.7.12}$$

Hence,

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 2. \tag{4.7.13}$$

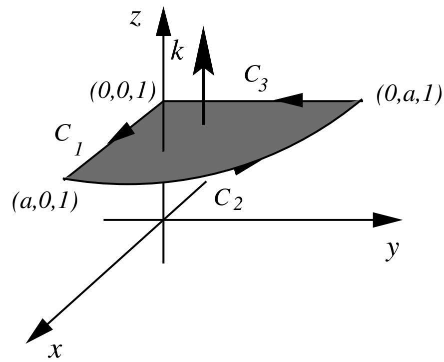


Figure 4.7.2: Diagram for the verification of Stokes' theorem in Example 4.7.2.

Turning to the other side of the equation,

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2 & 2x & z^2 \end{vmatrix} = 2\mathbf{k}. \tag{4.7.14}$$

Our line integral has been such that the normal vector must be $\mathbf{n} = \mathbf{k}$. Therefore,

$$\iint_S \nabla \times \mathbf{F} \cdot \mathbf{n}\, d\sigma = \int_0^1 \int_0^1 2\mathbf{k} \cdot \mathbf{k}\, dx\, dy = 2 \tag{4.7.15}$$

and Stokes' theorem is verified for this special case. □

• **Example 4.7.2**

Let us verify Stokes' theorem using the vector field $\mathbf{F} = (x^2 - y)\mathbf{i} + 4z\mathbf{j} + x^2\mathbf{k}$, where the closed contour consists of the x and y coordinate axes and that portion of the circle $x^2 + y^2 = a^2$ that lies in the first quadrant with $z = 1$. See Figure 4.7.2.

The line integral consists of three parts:

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r} + \int_{C_3} \mathbf{F} \cdot d\mathbf{r}. \tag{4.7.16}$$

Along C_1, x varies while $y = 0$ and $z = 1$. Therefore,

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_0^a (x^2 - 0)\, dx + 4 \cdot 1 \cdot 0 + x^2 \cdot 0 = \frac{a^3}{3}. \tag{4.7.17}$$

Along the circle C_2, we use polar coordinates with $x = a\cos(t)$, $y = a\sin(t)$, and $z = 1$. Therefore,

$$\int_{C_2} \mathbf{F} \cdot d\mathbf{r} = \int_0^{\pi/2} [a^2\cos^2(t) - a\sin(t)][-a\sin(t)\, dt] + 4 \cdot 1 \cdot a\cos(t)\, dt + a^2\cos^2(t) \cdot 0, \tag{4.7.18}$$

$$= \int_0^{\pi/2} -a^3\cos^2(t)\sin(t)\, dt + a^2\sin^2(t)\, dt + 4a\cos(t)\, dt \tag{4.7.19}$$

$$= \frac{a^3}{3}\cos^3(t)\Big|_0^{\pi/2} + \frac{a^2}{2}\left[t - \frac{1}{2}\sin(2t)\right]\Big|_0^{\pi/2} + 4a\sin(t)\Big|_0^{\pi/2} \tag{4.7.20}$$

$$= -\frac{a^3}{3} + \frac{a^2\pi}{4} + 4a, \tag{4.7.21}$$

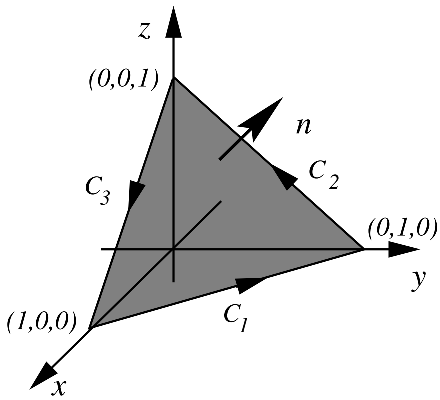


Figure 4.7.3: Diagram for the verification of Stokes' theorem in Example 4.7.3.

because $dx = -a\sin(t)\,dt$, and $dy = a\cos(t)\,dt$. Finally, along C_3, y varies with $x = 0$ and $z = 1$. Therefore,

$$\int_{C_3} \mathbf{F}\cdot d\mathbf{r} = \int_a^0 (0^2 - y)\cdot 0 + 4\cdot 1\cdot dy + 0^2\cdot 0 = -4a, \qquad (4.7.22)$$

so that

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \frac{a^2\pi}{4}. \qquad (4.7.23)$$

Turning to the other side of the equation,

$$\nabla\times\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2 - y & 4z & x^2 \end{vmatrix} = -4\mathbf{i} - 2x\mathbf{j} + \mathbf{k}. \qquad (4.7.24)$$

From the path of our line integral, our unit normal vector must be $\mathbf{n} = \mathbf{k}$. Then,

$$\iint_S \nabla\times\mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^a \int_0^{\pi/2} [-4\mathbf{i} - 2r\cos(\theta)\mathbf{j} + \mathbf{k}]\cdot\mathbf{k}\,r\,d\theta\,dr = \frac{\pi a^2}{4} \qquad (4.7.25)$$

and Stokes' theorem is verified for this case. □

• Example 4.7.3

Let us verify Stokes' theorem using the vector field $\mathbf{F} = 2yz\mathbf{i} - (x+3y-2)\mathbf{j} + (x^2+z)\mathbf{k}$, where the closed triangular region is that portion of the plane $x + y + z = 1$ that lies in the first octant.

As shown in Figure 4.7.3, the closed line integration consists of three line integrals:

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot d\mathbf{r} + \int_{C_2} \mathbf{F}\cdot d\mathbf{r} + \int_{C_3} \mathbf{F}\cdot d\mathbf{r}. \qquad (4.7.26)$$

Along C_1, $z = 0$ and $y = 1 - x$. Therefore, using x as the independent variable,

$$\int_{C_1} \mathbf{F}\cdot d\mathbf{r} = \int_1^0 2(1-x)\cdot 0\cdot dx - (x+3-3x-2)(-dx) + (x^2+0)\cdot 0 = -x^2\big|_1^0 + x\big|_1^0 = 0. \qquad (4.7.27)$$

Along C_2, $x=0$ and $y=1-z$. Thus,

$$\int_{C_2} \mathbf{F}\cdot d\mathbf{r} = \int_0^1 2(1-z)z\cdot 0 - (0+3-3z-2)(-dz) + (0^2+z)\,dz = -\left.\tfrac{3}{2}z^2 + z + \tfrac{1}{2}z^2\right|_0^1 = 0. \tag{4.7.28}$$

Finally, along C_3, $y=0$ and $z=1-x$. Hence,

$$\int_{C_3} \mathbf{F}\cdot d\mathbf{r} = \int_0^1 2\cdot 0\cdot(1-x)\,dx - (x+0-2)\cdot 0 + (x^2+1-x)(-dx) = -\left.\tfrac{1}{3}x^3 - x + \tfrac{1}{2}x^2\right|_0^1 = -\tfrac{5}{6}. \tag{4.7.29}$$

Thus,

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = -\tfrac{5}{6}. \tag{4.7.30}$$

On the other hand,

$$\nabla\times\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 2yz & -x-3y+2 & x^2+z \end{vmatrix} = (-2x+2y)\mathbf{j} + (-1-2z)\mathbf{k}. \tag{4.7.31}$$

To find $\mathbf{n}\,d\sigma$, we use the general coordinate system $x=u$, $y=v$, and $z=1-u-v$. Therefore, $\mathbf{r}=u\mathbf{i}+v\mathbf{j}+(1-u-v)\mathbf{k}$ and

$$\mathbf{r}_u\times\mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{vmatrix} = \mathbf{i}+\mathbf{j}+\mathbf{k}. \tag{4.7.32}$$

Thus,

$$\iint_S \nabla\times\mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^1\int_0^{1-u} [(-2u+2v)\mathbf{j} + (-1-2+2u+2v)\mathbf{k}]\cdot[\mathbf{i}+\mathbf{j}+\mathbf{k}]\,dv\,du \tag{4.7.33}$$

$$= \int_0^1\int_0^{1-u} (4v-3)\,dv\,du = \int_0^1 [2(1-u)^2 - 3(1-u)]\,du \tag{4.7.34}$$

$$= \int_0^1 (-1-u+2u^2)\,du = -\tfrac{5}{6} \tag{4.7.35}$$

and Stokes' theorem is verified for this case.

Problems

Verify Stokes' theorem using the following vector fields and surfaces:

1. $\mathbf{F}=5y\mathbf{i}-5x\mathbf{j}+3z\mathbf{k}$ and the surface S is that portion of the plane $z=1$ with the square at the vertices $(0,0,1)$, $(1,0,1)$, $(1,1,1)$, and $(0,1,1)$.

2. $\mathbf{F}=x^2\mathbf{i}+y^2\mathbf{j}+z^2\mathbf{k}$ and the surface S is the rectangular portion of the plane $z=2$ defined by the corners $(0,0,2)$, $(2,0,2)$, $(2,1,2)$, and $(0,1,2)$.

3. $\mathbf{F}=z\mathbf{i}+x\mathbf{j}+y\mathbf{k}$ and the surface S is the triangular portion of the plane $z=1$ defined by the vertices $(0,0,1)$, $(2,0,1)$, and $(0,2,1)$.

4. $\mathbf{F} = 2z\mathbf{i} - 3x\mathbf{j} + 4y\mathbf{k}$ and the surface S is that portion of the plane $z = 5$ within the cylinder $x^2 + y^2 = 4$.

5. $\mathbf{F} = z\mathbf{i} + x\mathbf{j} + y\mathbf{k}$ and the surface S is that portion of the plane $z = 3$ bounded by the lines $y = 0$, $x = 0$, and $x^2 + y^2 = 4$.

6. $\mathbf{F} = (2z + x)\mathbf{i} + (y - z)\mathbf{j} + (x + y)\mathbf{k}$ and the surface S is the interior of the triangularly shaped plane with vertices at $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$.

7. $\mathbf{F} = z\mathbf{i} + x\mathbf{j} + y\mathbf{k}$ and the surface S is that portion of the plane $2x + y + 2z = 6$ in the first octant.

8. $\mathbf{F} = x\mathbf{i} + xz\mathbf{j} + y\mathbf{k}$ and the surface S is that portion of the paraboloid $z = 9 - x^2 - y^2$ within the cylinder $x^2 + y^2 = 4$.

4.8 DIVERGENCE THEOREM

Although Stokes' theorem is useful in computing closed line integrals, it is usually very difficult to go the other way and convert a surface integral into a closed line integral because the integrand must have a very special form, namely $\nabla \times \mathbf{F} \cdot \mathbf{n}$. In this section we introduce a theorem that allows with equal facility the conversion of a closed surface integral into a volume integral and *vice versa*. Furthermore, if we can convert a given surface integral into a closed one by the introduction of a simple surface (for example, closing a hemispheric surface by adding an equatorial plate), it may be easier to use the divergence theorem and subtract off the contribution from the new surface integral rather than do the original problem.

This relationship between a closed surface integral and a volume integral involving the divergence operator is

The Divergence or Gauss's Theorem: *Let V be a closed and bounded region in three-dimensional space with a piece-wise smooth boundary S that is oriented outward. Let $\mathbf{F} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$ be a vector field for which P, Q, and R are continuous and have continuous first partial derivatives in a region of three-dimensional space containing V. Then*

$$\oint\!\!\!\!\oint_S \mathbf{F} \cdot \mathbf{n}\, d\sigma = \iiint_V \nabla \cdot \mathbf{F}\, dV. \qquad \textbf{(4.8.1)}$$

Here, the circle on the double integral signs denotes a closed surface integral.

A nonrigorous proof of Gauss's theorem is as follows. Imagine that our volume V is broken down into small elements $d\tau$ of volume of any shape so long as they include all of the original volume. In general, the surfaces of these elements are composed of common interfaces between adjoining elements. However, for the elements at the periphery of V, part of their surface will be part of the surface S that encloses V. Now $d\Phi = \nabla \cdot \mathbf{F}\, d\tau$ is the net flux of the vector $\mathbf{F}$ out from the element $d\tau$. At the common interface between elements, the flux *out* of one element equals the flux *into* its neighbor. Therefore, the sum of all such terms yields

$$\Phi = \iiint_V \nabla \cdot \mathbf{F}\, d\tau \qquad \textbf{(4.8.2)}$$

Carl Friedrich Gauss (1777–1855), the prince of mathematicians, must be on the list of the greatest mathematicians who ever lived. Gauss, a child prodigy, is almost as well known for what he did not publish during his lifetime as for what he did. This is true of Gauss's divergence theorem, which he proved while working on the theory of gravitation. It was only when his notebooks were published in 1898 that his precedence over the published work of Ostrogradsky (1801–1862) was established. (Portrait courtesy of Photo AKG, London, with permission.)

and all of the contributions from these common interfaces cancel; only the contribution from the parts on the outer surface S is left. These contributions, when added together, give $\oiint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma$ over S and the proof is completed. □

• Example 4.8.1

Let us verify the divergence theorem using the vector field $\mathbf{F} = 4x\mathbf{i} - 2y^2\mathbf{j} + z^2\mathbf{k}$ and the enclosed surface is the cylinder $x^2 + y^2 = 4$, $z = 0$, and $z = 3$. See Figure 4.8.1.

We begin by computing the volume integration. Because

$$\nabla\cdot\mathbf{F} = \frac{\partial(4x)}{\partial x} + \frac{\partial(-2y^2)}{\partial y} + \frac{\partial(z^2)}{\partial z} = 4 - 4y + 2z, \tag{4.8.3}$$

$$\iiint_V \nabla\cdot\mathbf{F}\,dV = \iiint_V (4 - 4y + 2z)\,dV \tag{4.8.4}$$

$$= \int_0^3\int_0^2\int_0^{2\pi} [4 - 4r\sin(\theta) + 2z]\,d\theta\,r\,dr\,dz \tag{4.8.5}$$

$$= \int_0^3\int_0^2 \left[4\theta\Big|_0^{2\pi} + 4r\cos(\theta)\Big|_0^{2\pi} + 2z\theta\Big|_0^{2\pi}\right] r\,dr\,dz \tag{4.8.6}$$

$$= \int_0^3\int_0^2 (8\pi + 4\pi z)\,r\,dr\,dz = \int_0^3 4\pi(2+z)\tfrac{1}{2}r^2\Big|_0^2\,dz \tag{4.8.7}$$

$$= 4\pi\int_0^3 2(2+z)\,dz = 8\pi(2z + \tfrac{1}{2}z^2)\Big|_0^3 = 84\pi. \tag{4.8.8}$$

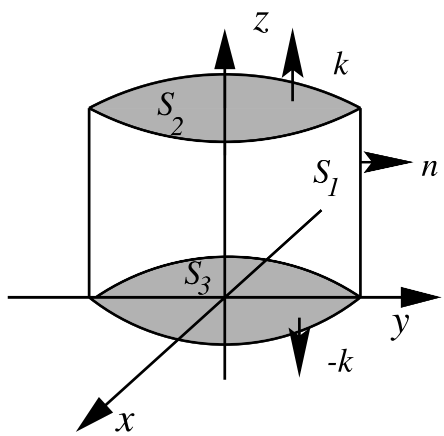


Figure 4.8.1: Diagram for the verification of the divergence theorem in Example 4.8.1.

Turning to the surface integration, we have three surfaces:

$$\oiint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \iint_{S_1}\mathbf{F}\cdot\mathbf{n}\,d\sigma + \iint_{S_2}\mathbf{F}\cdot\mathbf{n}\,d\sigma + \iint_{S_3}\mathbf{F}\cdot\mathbf{n}\,d\sigma. \qquad (\mathbf{4.8.9})$$

The first integral is over the exterior to the cylinder. Because the surface is defined by $f(x,y,z)=x^2+y^2=4$,

$$\mathbf{n}=\frac{\nabla f}{|\nabla f|}=\frac{2x\mathbf{i}+2y\mathbf{j}}{\sqrt{4x^2+4y^2}}=\frac{x}{2}\mathbf{i}+\frac{y}{2}\mathbf{j}. \qquad (\mathbf{4.8.10})$$

Therefore,

$$\iint_{S_1}\mathbf{F}\cdot\mathbf{n}\,d\sigma=\iint_{S_1}(2x^2-y^3)\,d\sigma=\int_0^3\int_0^{2\pi}\left\{2[2\cos(\theta)]^2-[2\sin(\theta)]^3\right\}\,2\,d\theta\,dz \qquad (\mathbf{4.8.11})$$

$$=8\int_0^3\int_0^{2\pi}\left\{\tfrac{1}{2}[1+\cos(2\theta)]-\sin(\theta)+\cos^2(\theta)\sin(\theta)\right\}2\,d\theta\,dz \qquad (\mathbf{4.8.12})$$

$$=16\int_0^3\left[\tfrac{1}{2}\theta+\tfrac{1}{4}\sin(2\theta)+\cos(\theta)-\tfrac{1}{3}\cos^3(\theta)\right]\Big|_0^{2\pi}\,dz \qquad (\mathbf{4.8.13})$$

$$=16\pi\int_0^3 dz=48\pi, \qquad (\mathbf{4.8.14})$$

because $x=2\cos(\theta)$, $y=2\sin(\theta)$, and $d\sigma=2\,d\theta\,dz$ in cylindrical coordinates.

Along the top of the cylinder, $z=3$, the outward pointing normal is $\mathbf{n}=\mathbf{k}$, and $d\sigma=r\,dr\,d\theta$. Then,

$$\iint_{S_2}\mathbf{F}\cdot\mathbf{n}\,d\sigma=\iint_{S_2}z^2\,d\sigma=\int_0^{2\pi}\int_0^2 9\,r\,dr\,d\theta=2\pi\times 9\times 2=36\pi. \qquad (\mathbf{4.8.15})$$

However, along the bottom of the cylinder, $z=0$, the outward pointing normal is $\mathbf{n}=-\mathbf{k}$ and $d\sigma=r\,dr\,d\theta$. Then,

$$\iint_{S_3}\mathbf{F}\cdot\mathbf{n}\,d\sigma=\iint_{S_3}z^2\,d\sigma=\int_0^{2\pi}\int_0^2 0\,r\,dr\,d\theta=0. \qquad (\mathbf{4.8.16})$$

Consequently, the flux out of the entire cylinder is

$$\oiint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = 48\pi + 36\pi + 0 = 84\pi, \tag{4.8.17}$$

and the divergence theorem is verified for this special case. □

• Example 4.8.2

Let us verify the divergence theorem given the vector field $\mathbf{F} = 3x^2y^2\mathbf{i} + y\mathbf{j} - 6xy^2z\mathbf{k}$ and the volume is the region bounded by the paraboloid $z = x^2 + y^2$, and the plane $z = 2y$. See Figure 4.8.2.

Computing the divergence,

$$\nabla\cdot\mathbf{F} = \frac{\partial(3x^2y^2)}{\partial x} + \frac{\partial(y)}{\partial y} + \frac{\partial(-6xy^2z)}{\partial z} = 6xy^2 + 1 - 6xy^2 = 1. \tag{4.8.18}$$

Then,

$$\iiint_V \nabla\cdot\mathbf{F}\,dV = \iiint_V dV = \int_0^{\pi}\int_0^{2\sin(\theta)}\int_{r^2}^{2r\sin(\theta)} dz\,r\,dr\,d\theta \tag{4.8.19}$$

$$= \int_0^{\pi}\int_0^{2\sin(\theta)} [2r\sin(\theta) - r^2]\,r\,dr\,d\theta \tag{4.8.20}$$

$$= \int_0^{\pi}\left[\tfrac{2}{3}r^3\Big|_0^{2\sin(\theta)}\sin(\theta) - \tfrac{1}{4}r^4\Big|_0^{2\sin(\theta)}\right] d\theta \tag{4.8.21}$$

$$= \int_0^{\pi}\left[\tfrac{16}{3}\sin^4(\theta) - 4\sin^4(\theta)\right] d\theta = \int_0^{\pi}\tfrac{4}{3}\sin^4(\theta)\,d\theta \tag{4.8.22}$$

$$= \tfrac{1}{3}\int_0^{\pi}[1 - 2\cos(2\theta) + \cos^2(2\theta)]\,d\theta \tag{4.8.23}$$

$$= \tfrac{1}{3}\left[\theta\Big|_0^{\pi} - \sin(2\theta)\Big|_0^{\pi} + \tfrac{1}{2}\theta\Big|_0^{\pi} + \tfrac{1}{8}\sin(4\theta)\Big|_0^{\pi}\right] = \tfrac{\pi}{2}. \tag{4.8.24}$$

The limits in the radial direction are given by the intersection of the paraboloid and plane: $r^2 = 2r\sin(\theta)$, or $r = 2\sin(\theta)$, and y is greater than zero.

Turning to the surface integration, we have two surfaces:

$$\oiint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \iint_{S_1} \mathbf{F}\cdot\mathbf{n}\,d\sigma + \iint_{S_2} \mathbf{F}\cdot\mathbf{n}\,d\sigma, \tag{4.8.25}$$

where S_1 is the plane $z = 2y$, and S_2 is the paraboloid. For either surface, polar coordinates are best so that $x = r\cos(\theta)$, and $y = r\sin(\theta)$. For the integration over the plane, $z = 2r\sin(\theta)$. Therefore,

$$\mathbf{r} = r\cos(\theta)\mathbf{i} + r\sin(\theta)\mathbf{j} + 2r\sin(\theta)\mathbf{k}, \tag{4.8.26}$$

so that

$$\mathbf{r}_r = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j} + 2\sin(\theta)\mathbf{k}, \tag{4.8.27}$$

and

$$\mathbf{r}_\theta = -r\sin(\theta)\mathbf{i} + r\cos(\theta)\mathbf{j} + 2r\cos(\theta)\mathbf{k}. \tag{4.8.28}$$

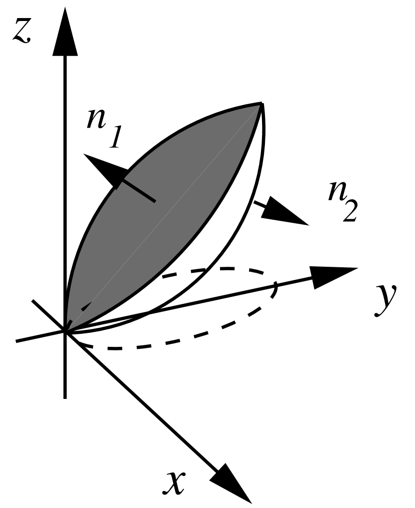


Figure 4.8.2: Diagram for the verification of the divergence theorem in Example 4.8.2. The dashed line denotes the curve $r = 2\sin(\theta)$.

Then,

$$\mathbf{r}_r \times \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos(\theta) & \sin(\theta) & 2\sin(\theta) \\ -r\sin(\theta) & r\cos(\theta) & 2r\cos(\theta) \end{vmatrix} = -2r\mathbf{j} + r\mathbf{k}. \tag{4.8.29}$$

This is an outwardly pointing normal so that we can immediately set up the surface integral:

$$\begin{aligned} \iint_{S_1} \mathbf{F}\cdot\mathbf{n}\,d\sigma &= \int_0^\pi \int_0^{2\sin(\theta)} \big\{3r^4\cos^2(\theta)\sin^2(\theta)\mathbf{i} + r\sin(\theta)\mathbf{j} \\ &\qquad - 6[2r\sin(\theta)][r\cos(\theta)][r^2\sin^2(\theta)]\mathbf{k}\big\}\cdot\big(-2r\mathbf{j} + r\mathbf{k}\big)\,dr\,d\theta && (4.8.30) \\ &= \int_0^\pi \int_0^{2\sin(\theta)} \left[-2r^2\sin(\theta) - 12r^5\sin^3(\theta)\cos(\theta)\right] dr\,d\theta && (4.8.31) \\ &= \int_0^\pi \left[-\tfrac{2}{3}r^3\Big|_0^{2\sin(\theta)}\sin(\theta) - 2r^6\Big|_0^{2\sin(\theta)}\sin^3(\theta)\cos(\theta)\right] d\theta && (4.8.32) \\ &= \int_0^\pi \left[-\tfrac{16}{3}\sin^4(\theta) - 128\sin^9(\theta)\cos(\theta)\right] d\theta && (4.8.33) \\ &= -\tfrac{4}{3}\left[\theta\Big|_0^\pi - \sin(2\theta)\Big|_0^\pi + \tfrac{1}{2}\theta\Big|_0^\pi + \tfrac{1}{8}\sin(4\theta)\Big|_0^\pi\right] - \tfrac{64}{5}\sin^{10}(\theta)\Big|_0^\pi && (4.8.34) \\ &= -2\pi. && (4.8.35) \end{aligned}$$

For the surface of the paraboloid,

$$\mathbf{r} = r\cos(\theta)\mathbf{i} + r\sin(\theta)\mathbf{j} + r^2\mathbf{k}, \tag{4.8.36}$$

so that

$$\mathbf{r}_r = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j} + 2r\mathbf{k}, \tag{4.8.37}$$

and

$$\mathbf{r}_\theta = -r\sin(\theta)\mathbf{i} + r\cos(\theta)\mathbf{j}. \tag{4.8.38}$$

Then,

$$\mathbf{r}_r \times \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos(\theta) & \sin(\theta) & 2r \\ -r\sin(\theta) & r\cos(\theta) & 0 \end{vmatrix} = -2r^2\cos(\theta)\mathbf{i} - 2r^2\sin(\theta)\mathbf{j} + r\mathbf{k}. \tag{4.8.39}$$

This is an inwardly pointing normal, so that we must take the negative of it before we do the surface integral. Then,

$$\iint_{S_2} \mathbf{F}\cdot\mathbf{n}\,d\sigma = \int_0^\pi \int_0^{2\sin(\theta)} \left\{3r^4\cos^2(\theta)\sin^2(\theta)\mathbf{i} + r\sin(\theta)\mathbf{j} - 6r^2[r\cos(\theta)][r^2\sin^2(\theta)]\mathbf{k}\right\} \cdot \left[2r^2\cos(\theta)\mathbf{i} + 2r^2\sin(\theta)\mathbf{j} - r\mathbf{k}\right] dr\,d\theta \tag{4.8.40}$$

$$= \int_0^\pi \int_0^{2\sin(\theta)} \left[6r^6\cos^3(\theta)\sin^2(\theta) + 2r^3\sin^2(\theta) + 6r^6\cos(\theta)\sin^2(\theta)\right] dr\,d\theta \tag{4.8.41}$$

$$= \int_0^\pi \left[\tfrac{6}{7}r^7\Big|_0^{2\sin(\theta)}\cos^3(\theta)\sin^2(\theta) + \tfrac{1}{2}r^4\Big|_0^{2\sin(\theta)}\sin^2(\theta) + \tfrac{6}{7}r^7\Big|_0^{2\sin(\theta)}\cos(\theta)\sin^2(\theta)\right] d\theta \tag{4.8.42}$$

$$= \int_0^\pi \left\{\tfrac{768}{7}\sin^9(\theta)[1-\sin^2(\theta)]\cos(\theta) + 8\sin^6(\theta) + \tfrac{768}{7}\sin^9(\theta)\cos(\theta)\right\} d\theta \tag{4.8.43}$$

$$= \tfrac{1536}{70}\sin^{10}(\theta)\Big|_0^\pi - \tfrac{64}{7}\sin^{12}(\theta)\Big|_0^\pi + \int_0^\pi [1-\cos(2\theta)]^3\,d\theta \tag{4.8.44}$$

$$= \int_0^\pi \left\{1 - 3\cos(2\theta) + 3\cos^2(2\theta) - \cos(2\theta)[1-\sin^2(2\theta)]\right\} d\theta \tag{4.8.45}$$

$$= \theta\Big|_0^\pi - \tfrac{3}{2}\sin(2\theta)\Big|_0^\pi + \tfrac{3}{2}[\theta + \tfrac{1}{4}\sin(4\theta)]\Big|_0^\pi - \tfrac{1}{2}\sin(2\theta)\Big|_0^\pi + \tfrac{1}{3}\sin^3(2\theta)\Big|_0^\pi \tag{4.8.46}$$

$$= \pi + \tfrac{3}{2}\pi = \tfrac{5}{2}\pi. \tag{4.8.47}$$

Consequently,

$$\oiint_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = -2\pi + \tfrac{5}{2}\pi = \tfrac{1}{2}\pi, \tag{4.8.48}$$

and the divergence theorem is verified for this special case. □

• Example 4.8.3: Archimedes' principle

Consider a solid[6] of volume V and surface S that is immersed in a vessel filled with a fluid of density ρ. The pressure field p in the fluid is a function of the distance from the liquid/air interface and equals

$$p = p_0 - \rho g z, \tag{4.8.49}$$

[6] Adapted from Altintas, A., 1990: Archimedes' principle as an application of the divergence theorem. *IEEE Trans. Educ.*, **33**, 222.

where g is the gravitational acceleration, z is the vertical distance measured from the interface (increasing in the $\mathbf{k}$ direction), and p_0 is the constant pressure along the liquid/air interface.

If we define $\mathbf{F} = -p\mathbf{k}$, then $\mathbf{F} \cdot \mathbf{n}\, d\sigma$ is the vertical component of the force on the surface due to the pressure and $\oiint_S \mathbf{F} \cdot \mathbf{n}\, d\sigma$ is the total lift. Using the divergence theorem and noting that $\nabla \cdot \mathbf{F} = \rho g$, the total lift also equals

$$\iiint_V \nabla \cdot \mathbf{F}\, dV = \rho g \iiint_V dV = \rho g V, \tag{4.8.50}$$

which is the weight of the displaced liquid. This is *Archimedes' principle*: The buoyant force on a solid immersed in a fluid of constant density equals the weight of the fluid displaced. □

• Example 4.8.4: Conservation of charge

Let a charge of density ρ flow with an average velocity $\mathbf{v}$. Then the charge crossing the element $d\mathbf{S}$ per unit time is $\rho\mathbf{v} \cdot d\mathbf{S} = \mathbf{J} \cdot d\mathbf{S}$, where $\mathbf{J}$ is defined as the conduction current vector or current density vector. The current across any surface drawn in the medium is $\oiint_S \mathbf{J} \cdot d\mathbf{S}$.

The total charge inside the closed surface is $\iiint_V \rho\, dV$. If there are no sources or sinks inside the surface, the rate at which the charge decreases is $-\iiint_V \rho_t\, dV$. Because this change is due to the outward flow of charge,

$$-\iiint_V \frac{\partial \rho}{\partial t}\, dV = \oiint_S \mathbf{J} \cdot d\mathbf{S}. \tag{4.8.51}$$

Applying the divergence theorem,

$$\iiint_V \left(\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} \right) dV = 0. \tag{4.8.52}$$

Because the result holds true for any arbitrary volume, the integrand must vanish identically and we have the equation of continuity or the *equation of conservation of charge*:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \tag{4.8.53}$$

Problems

Verify the divergence theorem using the following vector fields and volumes:

1. $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$ and the volume V is the cube cut from the first octant by the planes $x = 1$, $y = 1$, and $z = 1$.

2. $\mathbf{F} = xy\mathbf{i} + yz\mathbf{j} + xz\mathbf{k}$ and the volume V is the cube bounded by $0 \le x \le 1$, $0 \le y \le 1$, and $0 \le z \le 1$.

3. $\mathbf{F} = (y - x)\mathbf{i} + (z - y)\mathbf{j} + (y - x)\mathbf{k}$ and the volume V is the cube bounded by $-1 \le x \le 1$, $-1 \le y \le 1$, and $-1 \le z \le 1$.

4. $\mathbf{F} = x^2\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and the volume V is the cylinder defined by the surfaces $x^2 + y^2 = 1$, $z = 0$, and $z = 1$.

5. $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$ and the volume V is the cylinder defined by the surfaces $x^2 + y^2 = 4$, $z = 0$, and $z = 1$.

6. $\mathbf{F} = y^2\mathbf{i} + xz^3\mathbf{j} + (z-1)^2\mathbf{k}$ and the volume V is the cylinder bounded by the surface $x^2 + y^2 = 4$, and the planes $z = 1$ and $z = 5$.

7. $\mathbf{F} = 6xy\mathbf{i} + 4yz\mathbf{j} + xe^{-y}\mathbf{k}$ and the volume V is that region created by the plane $x+y+z = 1$, and the three coordinate planes.

8. $\mathbf{F} = y\mathbf{i} + xy\mathbf{j} - z\mathbf{k}$ and the volume V is that solid created by the paraboloid $z = x^2 + y^2$ and plane $z = 1$.

Further Readings

Davis, H. F., and A. D. Snider, 1995: *Introduction to Vector Analysis.* Wm. C. Brown Publ., 416 pp. Designed as a reference book for engineering majors.

Kendall, P. C., and D. E. Bourne, 1992: *Vector Analysis and Cartesian Tensors.* Wm. C. Brown, Publ., 304 pp. A clear introduction to the concepts and techniques of vector analysis.

Matthews, P. C., 2005: *Vector Calculus.* Springer, 200 pp. A good book for self-study with complete solutions to the problems.

Schey, H. M., 2005: *Div, Grad, Curl, and All That.* Chapman & Hall, 176 pp. A book to hone your vector calculus skills.

Chapter 5
Fourier Series

Fourier series arose during the eighteenth century as a formal solution to the classic wave equation. Later on, it was used to describe physical processes in which events recur in a regular pattern. For example, a musical note usually consists of a simple note, called the fundamental, and a series of auxiliary vibrations, called overtones. Fourier's theorem provides the mathematical language that allows us to precisely describe this complex structure.

5.1 FOURIER SERIES

One of the crowning glories[1] of nineteenth-century mathematics was the discovery that the infinite series

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \qquad (\mathbf{5.1.1})$$

can represent a function $f(t)$ under certain general conditions. This series, called a *Fourier series*, converges to the value of the function $f(t)$ at every point in the interval $[-L, L]$ with the possible exceptions of the points at any discontinuities and the endpoints of the interval.

[1] "Fourier's Theorem ... is not only one of the most beautiful results of modern analysis, but may be said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern physics. To mention only sonorous vibrations, the propagation of electric signals along a telegraph wire, and the conduction of heat by the earth's crust, as subjects in their generality intractable without it, is to give but a feeble idea of its importance." (Quote taken from Thomson, W., and P. G. Tait, 1879: *Treatise on Natural Philosophy, Part 1*. Cambridge University Press, Section 75.)

Because each term has a period of $2L$, the sum of the series also has the same period. The *fundamental* of the periodic function $f(t)$ is the $n = 1$ term while the *harmonics* are the remaining terms whose frequencies are integer multiples of the fundamental. We must now find some easy method for computing the coefficients a_n and b_n for a given function $f(t)$. As a first attempt, we integrate Equation 5.1.1 term by term[2] from $-L$ to L. On the right side, all of the integrals multiplied by a_n and b_n vanish because the average of $\cos(n\pi t/L)$ and $\sin(n\pi t/L)$ is zero. Therefore, we are left with

$$\boxed{a_0 = \frac{1}{L}\int_{-L}^{L} f(t)\,dt.} \tag{5.1.2}$$

Consequently a_0 is twice the mean value of $f(t)$ over one period.

We next multiply each side of Equation 5.1.1 by $\cos(m\pi t/L)$, where m is a fixed integer. Integrating from $-L$ to L,

$$\int_{-L}^{L} f(t)\cos\left(\frac{m\pi t}{L}\right)\,dt = \frac{a_0}{2}\int_{-L}^{L}\cos\left(\frac{m\pi t}{L}\right)\,dt + \sum_{n=1}^{\infty} a_n \int_{-L}^{L}\cos\left(\frac{n\pi t}{L}\right)\cos\left(\frac{m\pi t}{L}\right)\,dt$$
$$+ \sum_{n=1}^{\infty} b_n \int_{-L}^{L}\sin\left(\frac{n\pi t}{L}\right)\cos\left(\frac{m\pi t}{L}\right)\,dt. \tag{5.1.3}$$

The a_0 and b_n terms vanish by direct integration. Finally, all of the a_n integrals vanish when $n \neq m$. Consequently, Equation 5.1.3 simplifies to

$$\boxed{a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{n\pi t}{L}\right)\,dt,} \tag{5.1.4}$$

because $\int_{-L}^{L}\cos^2(n\pi t/L)\,dt = L$. Finally, by multiplying both sides of Equation 5.1.1 by $\sin(m\pi t/L)$ (m is again a fixed integer) and integrating from $-L$ to L,

$$\boxed{b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi t}{L}\right)\,dt.} \tag{5.1.5}$$

Although Equation 5.1.2, Equation 5.1.4, and Equation 5.1.5 give us a_0, a_n, and b_n for periodic functions over the interval $[-L, L]$, in certain situations it is convenient to use the interval $[\tau, \tau + 2L]$, where τ is any real number. In that case, Equation 5.1.1 still gives the

[2] We assume that the integration of the series can be carried out term by term. This is sometimes difficult to justify but we do it anyway.

Fourier series of $f(t)$ and

$$a_0 = \frac{1}{L}\int_{\tau}^{\tau+2L} f(t)\,dt,$$
$$a_n = \frac{1}{L}\int_{\tau}^{\tau+2L} f(t)\cos\left(\frac{n\pi t}{L}\right)\,dt, \quad \textbf{(5.1.6)}$$
$$b_n = \frac{1}{L}\int_{\tau}^{\tau+2L} f(t)\sin\left(\frac{n\pi t}{L}\right)\,dt.$$

These results follow when we recall that the function $f(t)$ is a periodic function that extends from minus infinity to plus infinity. The results must remain unchanged, therefore, when we shift from the interval $[-L, L]$ to the new interval $[\tau, \tau + 2L]$.

We now ask the question: what types of functions have Fourier series? Secondly, if a function is discontinuous at a point, what value will the Fourier series give? Dirichlet[3,4] answered these questions in the first half of the nineteenth century. His results may be summarized as follows.

Dirichlet's Theorem: *If for the interval $[-L, L]$ the function $f(t)$ (1) is single-valued, (2) is bounded, (3) has at most a finite number of maxima and minima, and (4) has only a finite number of discontinuities (piecewise continuous), and if (5) $f(t + 2L) = f(t)$ for values of t outside of $[-L, L]$, then*

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \quad (\textbf{5.1.7})$$

converges to $f(t)$ as $N \to \infty$ at values of t for which $f(t)$ is continuous and to $\frac{1}{2}[f(t^+) + f(t^-)]$ at points of discontinuity. The quantities t^+ and t^- denote points infinitesimally to the right and left of t. The coefficients in Equation 5.1.7 are given by Equation 5.1.2, Equation 5.1.4, and Equation 5.1.5. A function $f(t)$ is bounded if the inequality $|f(t)| \le M$ holds for some constant M for all values of t. Because the *Dirichlet's conditions* (1)–(4) are very mild, it is very rare that a convergent Fourier series does not exist for a function that appears in an engineering or scientific problem. □

• Example 5.1.1

Let us find the Fourier series for the function

$$f(t) = \begin{cases} 0, & -\pi < t \le 0, \\ t, & 0 \le t < \pi. \end{cases} \quad (\textbf{5.1.8})$$

We compute the Fourier coefficients a_n and b_n using Equation 5.1.6 by letting $L = \pi$ and $\tau = -\pi$. We then find that

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,dt = \frac{1}{\pi}\int_0^{\pi} t\,dt = \frac{\pi}{2}, \quad (\textbf{5.1.9})$$

[3] Dirichlet, P. G. L., 1829: Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données. *J. Reine Angew. Math.*, **4**, 157–169.

[4] Dirichlet, P. G. L., 1837: Sur l'usage des intégrales définies dans la sommation des séries finies ou infinies. *J. Reine Angew. Math.*, **17**, 57–67.

A product of the French Revolution, (Jean Baptiste) Joseph Fourier (1768–1830) held positions within the Napoleonic Empire during his early career. After Napoleon's fall from power, Fourier devoted his talents exclusively to science. Although he won the Institut de France prize in 1811 for his work on heat diffusion, criticism of its mathematical rigor and generality led him to publish the classic book *Théorie analytique de la chaleur* in 1823. Within this book he introduced the world to the series that bears his name. (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

$$a_n = \frac{1}{\pi}\int_0^\pi t\cos(nt)\,dt = \frac{1}{\pi}\left[\frac{t\sin(nt)}{n} + \frac{\cos(nt)}{n^2}\right]\Bigg|_0^\pi = \frac{\cos(n\pi)-1}{n^2\pi} = \frac{(-1)^n-1}{n^2\pi} \tag{5.1.10}$$

because $\cos(n\pi) = (-1)^n$, and

$$b_n = \frac{1}{\pi}\int_0^\pi t\sin(nt)\,dt = \frac{1}{\pi}\left[\frac{-t\cos(nt)}{n} + \frac{\sin(nt)}{n^2}\right]\Bigg|_0^\pi = -\frac{\cos(n\pi)}{n} = \frac{(-1)^{n+1}}{n} \tag{5.1.11}$$

for $n = 1, 2, 3, \ldots$. Thus, the Fourier series for $f(t)$ is

$$f(t) = \frac{\pi}{4} + \sum_{n=1}^\infty \frac{(-1)^n-1}{n^2\pi}\cos(nt) + \frac{(-1)^{n+1}}{n}\sin(nt) \tag{5.1.12}$$

$$= \frac{\pi}{4} - \frac{2}{\pi}\sum_{m=1}^\infty \frac{\cos[(2m-1)t]}{(2m-1)^2} - \sum_{n=1}^\infty \frac{(-1)^n}{n}\sin(nt). \tag{5.1.13}$$

We note that at the points $t = \pm(2n-1)\pi$, where $n = 1, 2, 3, \ldots$, the function jumps from zero to π. To what value does the Fourier series converge at these points? From Dirichlet's theorem, the series converges to the average of the values of the function just to the right and left of the point of discontinuity, i.e., $(\pi + 0)/2 = \pi/2$. At the remaining points the series converges to $f(t)$.

Second to Gauss, Peter Gustav Lejeune Dirichlet (1805–1859) was Germany's leading mathematician during the first half of the nineteenth century. Initially drawn to number theory, his later studies in analysis and applied mathematics led him to consider the convergence of Fourier series. These studies eventually produced the modern concept of a function as a correspondence that associates with each real x in an interval some unique value denoted by $f(x)$. (Taken from the frontispiece of Dirichlet, P. G. L., 1889: *Werke.* Druck und Verlag von Georg Reimer, 644 pp.)

Figure 5.1.1 shows how well Equation 5.1.12 approximates the function by graphing various partial sums of this expansion as we include more and more terms (harmonics). The MATLAB script that created this figure is:

```
clear;
t = [-4:0.1:4]; % create time points in plot
f = zeros(size(t)); % initialize function f(t)
for k = 1:length(t) % construct function f(t)
  if t(k) < 0; f(k) = 0; else f(k) = t(k); end;
  if t(k) < -pi; f(k) = t(k) + 2*pi; end;
  if t(k) > pi ; f(k) = 0; end;
end
% initialize Fourier series with the mean term
fs = (pi/4) * ones(size(t));
clf % clear any figures
for n = 1:6
% create plot of truncated FS with only n harmonic
  fs = fs - (2/pi) * cos((2*n-1)*t) / (2*n-1)^2;
  fs = fs - (-1)^n * sin(n*t) / n;
  subplot(3,2,n), plot(t,fs,t,f,'--')
```

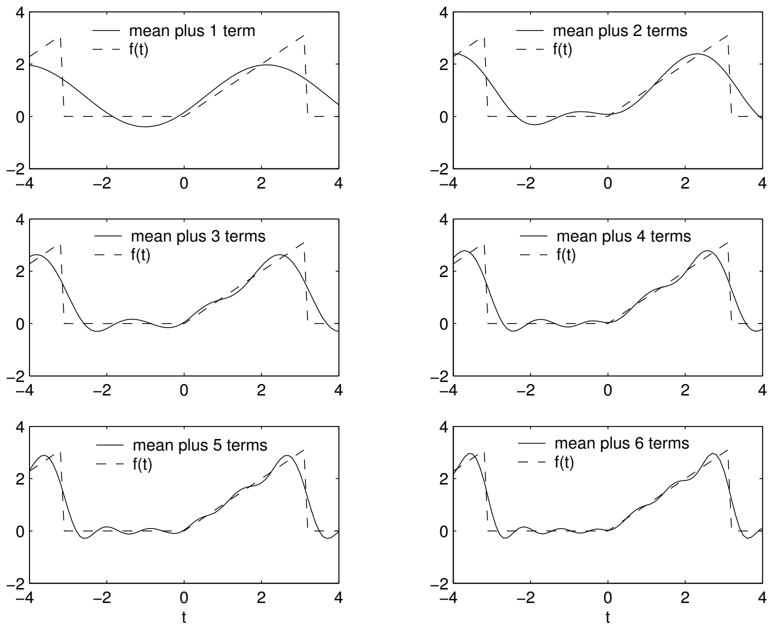


Figure 5.1.1: Partial sum of the Fourier series for Equation 5.1.8.

```
  if n==1
    legend('mean plus 1 term','f(t)'); legend boxoff;
  else
    legend(['mean plus ',num2str(n),' terms'],'f(t)')
    legend boxoff
  end
  if n >= 5; xlabel('t'); end;
end
```

As the figure shows, successive corrections are made to the mean value of the series, $\pi/2$. As each harmonic is added, the Fourier series fits the function better in the sense of least squares:

$$\int_{\tau}^{\tau+2L} [f(x) - f_N(x)]^2 \, dx = \text{ minimum}, \tag{5.1.14}$$

where $f_N(x)$ is the truncated Fourier series of N terms. □

• Example 5.1.2

Let us calculate the Fourier series of the function $f(t) = |t|$, which is defined over the range $-\pi \leq t \leq \pi$.

From the definition of the Fourier coefficients,

$$a_0 = \frac{1}{\pi} \left[\int_{-\pi}^{0} -t \, dt + \int_{0}^{\pi} t \, dt \right] = \frac{\pi}{2} + \frac{\pi}{2} = \pi, \tag{5.1.15}$$

$$a_n = \frac{1}{\pi} \left[\int_{-\pi}^{0} -t \cos(nt) \, dt + \int_{0}^{\pi} t \cos(nt) \, dt \right] \tag{5.1.16}$$

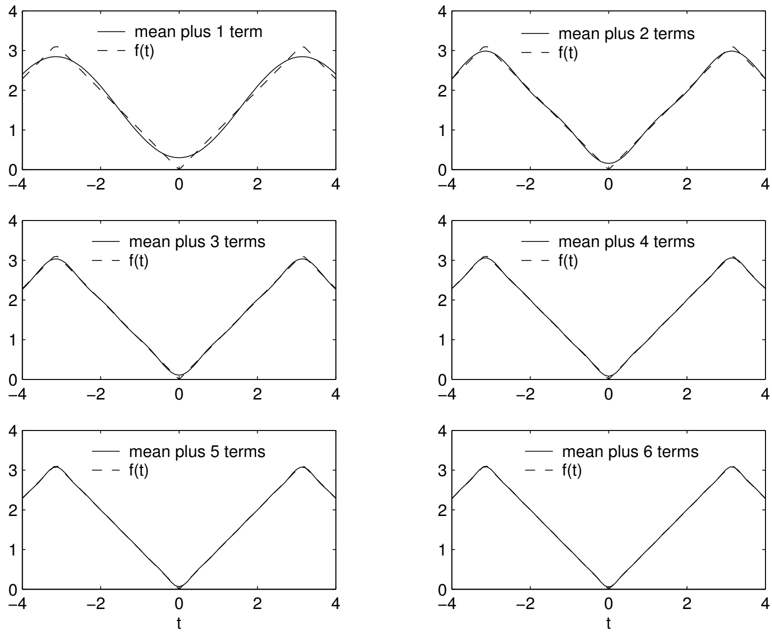


Figure 5.1.2: Partial sum of the Fourier series for $f(t) = |t|$.

$$= -\left.\frac{nt\sin(nt) + \cos(nt)}{n^2\pi}\right|_{-\pi}^{0} + \left.\frac{nt\sin(nt) + \cos(nt)}{n^2\pi}\right|_{0}^{\pi} \tag{5.1.17}$$

$$= \frac{2}{n^2\pi}[(-1)^n - 1] \tag{5.1.18}$$

and

$$b_n = \frac{1}{\pi}\left[\int_{-\pi}^{0} -t\sin(nt)\,dt + \int_{0}^{\pi} t\sin(nt)\,dt\right] \tag{5.1.19}$$

$$= \left.\frac{nt\cos(nt) - \sin(nt)}{n^2\pi}\right|_{-\pi}^{0} - \left.\frac{nt\cos(nt) - \sin(nt)}{n^2\pi}\right|_{0}^{\pi} = 0 \tag{5.1.20}$$

for $n = 1, 2, 3, \ldots$. Therefore,

$$|t| = \frac{\pi}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{[(-1)^n - 1]}{n^2}\cos(nt) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{m=1}^{\infty}\frac{\cos[(2m-1)t]}{(2m-1)^2} \tag{5.1.21}$$

for $-\pi \le t \le \pi$.

In Figure 5.1.2 we show how well Equation 5.1.21 approximates the function by graphing various partial sums of this expansion. As the figure shows, the Fourier series does very well even when we use very few terms. The reason for this rapid convergence is the nature of the function: it does not possess any jump discontinuities. □

• **Example 5.1.3**

Sometimes the function $f(t)$ is an even or odd function.[5] Can we use this property to simplify our work? The answer is yes.

[5] An even function $f_e(t)$ has the property that $f_e(-t) = f_e(t)$; an odd function $f_o(t)$ has the property that $f_o(-t) = -f_o(t)$.

Let $f(t)$ be an even function. Then

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(t)\,dt = \frac{2}{L}\int_0^L f(t)\,dt, \tag{5.1.22}$$

and

$$a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{n\pi t}{L}\right)\,dt = \frac{2}{L}\int_0^L f(t)\cos\left(\frac{n\pi t}{L}\right)\,dt, \tag{5.1.23}$$

whereas

$$b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi t}{L}\right)\,dt = 0. \tag{5.1.24}$$

Here we used the properties that $\int_{-L}^{L} f_e(x)\,dx = 2\int_0^L f_e(x)\,dx$ and $\int_{-L}^{L} f_o(x)dx = 0$. Thus, if we have an even function, we merely compute a_0 and a_n via Equation 5.1.22 and Equation 5.1.23, and $b_n = 0$. Because the corresponding series contains only cosine terms, it is often called a *Fourier cosine series.*

Similarly, if $f(t)$ is odd, then

$$a_0 = a_n = 0, \qquad \text{and} \qquad b_n = \frac{2}{L}\int_0^L f(t)\sin\left(\frac{n\pi t}{L}\right)\,dt. \tag{5.1.25}$$

Thus, if we have an odd function, we merely compute b_n via Equation 5.1.25 and $a_0 = a_n = 0$. Because the corresponding series contains only sine terms, it is often called a *Fourier sine series.* □

• Example 5.1.4

In the case when $f(x)$ consists of a constant and/or trigonometric functions, it is much easier to find the corresponding Fourier series by inspection rather than by using Equation 5.1.6. For example, let us find the Fourier series for $f(x) = \sin^2(x)$ defined over the range $-\pi \le x \le \pi$.

We begin by rewriting $f(x) = \sin^2(x)$ as $f(x) = \frac{1}{2}[1 - \cos(2x)]$. Next, we note that any function defined over the range $-\pi < x < \pi$ has the Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos(nx) + b_n\sin(nx) \tag{5.1.26}$$

$$= \frac{a_0}{2} + a_1\cos(x) + b_1\sin(x) + a_2\cos(2x) + b_2\sin(2x) + \cdots. \tag{5.1.27}$$

On the other hand,

$$f(x) = \tfrac{1}{2} - \tfrac{1}{2}\cos(2x) = \tfrac{1}{2} + 0\cos(x) + 0\sin(x) - \tfrac{1}{2}\cos(2x) + 0\sin(2x) + \cdots. \tag{5.1.28}$$

Consequently, by inspection, we can immediately write that

$$a_0 = 1, \quad a_1 = b_1 = 0, \quad a_2 = -\tfrac{1}{2}, \quad b_2 = 0, \quad a_n = b_n = 0, \quad n \ge 3. \tag{5.1.29}$$

Thus, instead of the usual expansion involving an infinite number of sine and cosine terms, our Fourier series contains only two terms and is simply

$$f(x) = \tfrac{1}{2} - \tfrac{1}{2}\cos(2x), \qquad -\pi \le x \le \pi. \tag{5.1.30}$$

□

• Example 5.1.5: Quieting snow tires

An application of Fourier series to a problem in industry occurred several years ago, when drivers found that snow tires produced a loud whine[6] on dry pavement. Tire sounds are produced primarily by the dynamic interaction of the tread elements with the road surface.[7] As each tread element passes through the contact patch, it contributes a pulse of acoustic energy to the total sound field radiated by the tire.

For evenly spaced treads we envision that the release of acoustic energy resembles the top of Figure 5.1.3. If we perform a Fourier analysis of this distribution, we find that

$$a_0 = \frac{1}{\pi}\left[\int_{-\pi/2-\epsilon}^{-\pi/2+\epsilon} 1\,dt + \int_{\pi/2-\epsilon}^{\pi/2+\epsilon} 1\,dt\right] = \frac{4\epsilon}{\pi}, \tag{5.1.31}$$

where ϵ is half of the width of the tread and

$$a_n = \frac{1}{\pi}\left[\int_{-\pi/2-\epsilon}^{-\pi/2+\epsilon} \cos(nt)\,dt + \int_{\pi/2-\epsilon}^{\pi/2+\epsilon} \cos(nt)\,dt\right] \tag{5.1.32}$$

$$= \frac{1}{n\pi}\left[\sin(nt)\Big|_{-\pi/2-\epsilon}^{-\pi/2+\epsilon} + \sin(nt)\Big|_{\pi/2-\epsilon}^{\pi/2+\epsilon}\right] \tag{5.1.33}$$

$$= \frac{1}{n\pi}\left[\sin\left(-\frac{n\pi}{2}+n\epsilon\right) - \sin\left(-\frac{n\pi}{2}-n\epsilon\right) + \sin\left(\frac{n\pi}{2}+n\epsilon\right) - \sin\left(\frac{n\pi}{2}-n\epsilon\right)\right] \tag{5.1.34}$$

$$= \frac{1}{n\pi}\left[2\cos\left(-\frac{n\pi}{2}\right) + 2\cos\left(\frac{n\pi}{2}\right)\right]\sin(n\epsilon) = \frac{4}{n\pi}\cos\left(\frac{n\pi}{2}\right)\sin(n\epsilon). \tag{5.1.35}$$

Because $f(t)$ is an even function, $b_n = 0$.

The question now arises of how to best illustrate our Fourier coefficients. In Section 5.4 we will show that any harmonic can be represented as a single wave $A_n\cos(n\pi t/L + \varphi_n)$ or $A_n\sin(n\pi t/L + \psi_n)$, where the amplitude $A_n = \sqrt{a_n^2 + b_n^2}$. In the bottom frame of Figure 5.1.3, MATLAB was used to plot this amplitude, usually called the *amplitude* or *frequency spectrum* $\frac{1}{2}\sqrt{a_n^2 + b_n^2}$, as a function of n for an arbitrarily chosen $\epsilon = \pi/12$. Although the value of ϵ will affect the exact shape of the spectrum, the qualitative arguments that we will present remain unchanged. We have added the factor $\frac{1}{2}$ so that our definition of the frequency spectrum is consistent with that for a complex Fourier series stated after Equation 5.5.12. The amplitude spectrum in Figure 5.1.3 shows that the spectrum for periodically placed tire treads has its largest amplitude at small n. This produces one loud tone plus strong harmonic overtones because the fundamental and its overtones are the dominant terms in the Fourier series representation.

Clearly this loud, monotone whine is undesirable. How might we avoid it? Just as soldiers marching in step produce a loud uniform sound, we suspect that our uniform tread pattern is the problem. Therefore, let us now vary the interval between the treads so that the distance between any tread and its nearest neighbor is not equal, as illustrated in Figure 5.1.4. Again we perform its Fourier analysis and obtain that

$$a_0 = \frac{1}{\pi}\left[\int_{-\pi/2-\epsilon}^{-\pi/2+\epsilon} 1\,dt + \int_{\pi/4-\epsilon}^{\pi/4+\epsilon} 1\,dt\right] = \frac{4\epsilon}{\pi}, \tag{5.1.36}$$

---

[6] See Varterasian, J. H., 1969: Math quiets rotating machines. *SAE J.*, **77(10)**, 53.

[7] Willett, P. R., 1975: Tire tread pattern sound generation. *Tire Sci. Tech.*, **3**, 252–266.

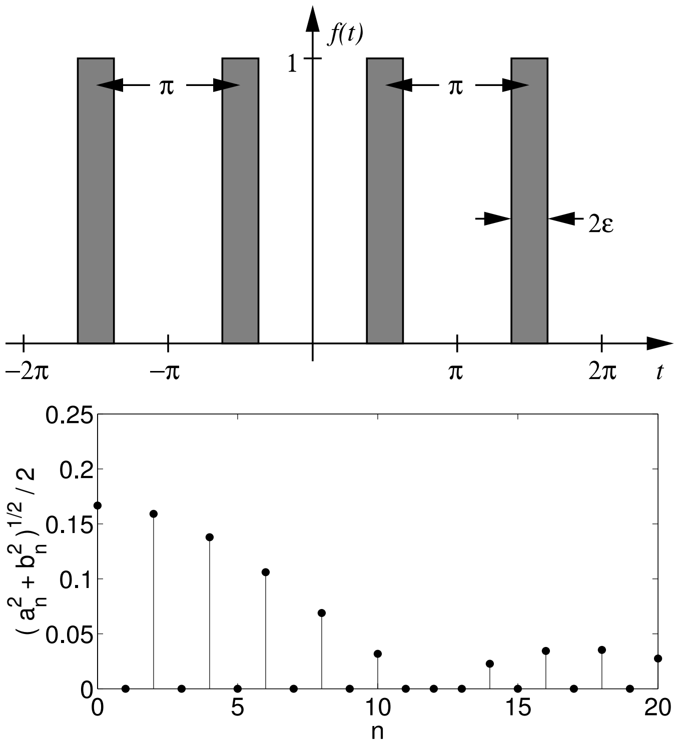


Figure 5.1.3: Temporal spacing (over two periods) and frequency spectrum of uniformly spaced snow tire treads.

$$a_n = \frac{1}{\pi}\left[\int_{-\pi/2-\epsilon}^{-\pi/2+\epsilon} \cos(nt)\,dt + \int_{\pi/4-\epsilon}^{\pi/4+\epsilon} \cos(nt)\,dt\right] \tag{5.1.37}$$

$$= \frac{1}{n\pi}\sin(nt)\bigg|_{-\pi/2-\epsilon}^{-\pi/2+\epsilon} + \frac{1}{n\pi}\sin(nt)\bigg|_{\pi/4-\epsilon}^{\pi/4+\epsilon} \tag{5.1.38}$$

$$= -\frac{1}{n\pi}\left[\sin\left(\frac{n\pi}{2}-n\epsilon\right) - \sin\left(\frac{n\pi}{2}+n\epsilon\right)\right] + \frac{1}{n\pi}\left[\sin\left(\frac{n\pi}{4}+n\epsilon\right) - \sin\left(\frac{n\pi}{4}-n\epsilon\right)\right] \tag{5.1.39}$$

$$a_n = \frac{2}{n\pi}\left[\cos\left(\frac{n\pi}{2}\right) + \cos\left(\frac{n\pi}{4}\right)\right]\sin(n\epsilon), \tag{5.1.40}$$

and

$$b_n = \frac{1}{\pi}\left[\int_{-\pi/2-\epsilon}^{-\pi/2+\epsilon} \sin(nt)\,dt + \int_{\pi/4-\epsilon}^{\pi/4+\epsilon} \sin(nt)\,dt\right] \tag{5.1.41}$$

$$= -\frac{1}{n\pi}\left[\cos\left(\frac{n\pi}{2}-n\epsilon\right) - \cos\left(\frac{n\pi}{2}+n\epsilon\right)\right] - \frac{1}{n\pi}\left[\cos\left(\frac{n\pi}{4}+n\epsilon\right) - \cos\left(\frac{n\pi}{4}-n\epsilon\right)\right] \tag{5.1.42}$$

$$= \frac{2}{n\pi}\left[\sin\left(\frac{n\pi}{4}\right) - \sin\left(\frac{n\pi}{2}\right)\right]\sin(n\epsilon). \tag{5.1.43}$$

The MATLAB script

```
epsilon = pi/12; % set up parameter for fs coefficient
n = 1:20; % number of harmonics
```

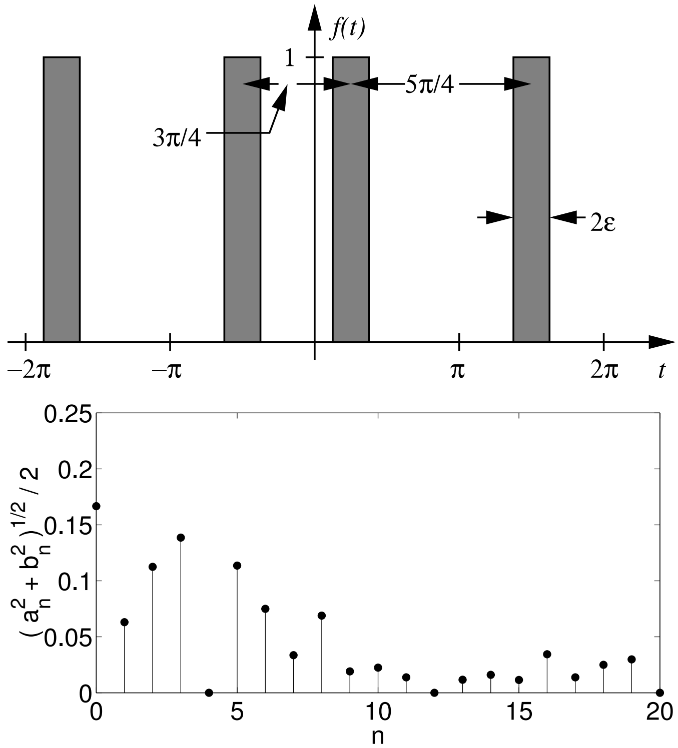


Figure 5.1.4: Temporal spacing and frequency spectrum of nonuniformly spaced snow tire treads.

```
arg1 = (pi/2)*n; arg2 = (pi/4)*n; arg3 = epsilon*n;
% compute the Fourier coefficient a_n
an = (cos(arg1) + cos(arg2)).*sin(arg3);
an = (2/pi) * an./n;
% compute the Fourier coefficient b_n
bn = (sin(arg2) - sin(arg1)).*sin(arg3);
bn = (2/pi) * bn./n;
% compute the magnitude
cn = 0.5 * sqrt(an.*an + bn.*bn);
% add in the a_0 term
cn = [2*epsilon/pi,cn];
n = [0,n];
clf % clear any figures
axes('FontSize',20) % set font size
stem(n,cn,'filled') % plot spectrum
set (gca,'PlotBoxAspectRatio',[8 4 1]) % set aspect ratio
xlabel('n') % label x-axis
ylabel('( a_n^2 + b_n^2 )^{1/2}/2') % label y-axis,
```

was used to compute the amplitude of each harmonic as a function of n and the results were plotted. See Figure 5.1.4. The important point is that our new choice for the spacing of the treads has reduced or eliminated some of the harmonics compared to the case of equally spaced treads. On the negative side we have excited some of the harmonics that were previously absent. However, the net effect is advantageous because the treads produce less noise at more frequencies rather than a lot of noise at a few select frequencies.

If we were to extend this technique so that the treads occurred at completely random positions, then the treads would produce very little noise at many frequencies and the total noise would be comparable to that generated by other sources within the car. To find the distribution of treads with the whitest noise[8] is a process of trial and error. Assuming a distribution, we can perform a Fourier analysis to obtain its frequency spectrum. If annoying peaks are present in the spectrum, we can then adjust the elements in the tread distribution that may contribute to the peak and analyze the revised distribution. You are finished when no peaks appear.

Problems

Find the Fourier series for the following functions. Using MATLAB, plot the Fourier spectrum. Then plot various partial sums and compare them against the exact function.

1. $f(t) = \begin{cases} 1, & -\pi < t < 0 \\ 0, & 0 < t < \pi \end{cases}$

2. $f(t) = \begin{cases} t, & -\pi < t \le 0 \\ 0, & 0 \le t < \pi \end{cases}$

3. $f(t) = \begin{cases} -\pi, & -\pi < t < 0 \\ t, & 0 < t < \pi \end{cases}$

4. $f(t) = \begin{cases} 1/2 + t, & -1 \le t \le 0 \\ 1/2 - t, & 0 \le t \le 1 \end{cases}$

5. $f(t) = \begin{cases} 0, & -\pi \le t \le 0 \\ t, & 0 \le t \le \pi/2 \\ \pi - t, & \pi/2 \le t \le \pi \end{cases}$

6. $f(t) = \begin{cases} 0, & -\pi \le t \le -\pi/2 \\ \sin(2t), & -\pi/2 \le t \le \pi/2 \\ 0, & \pi/2 \le t \le \pi \end{cases}$

7. $f(t) = e^{at}, \qquad -L < t < L$

8. $f(t) = t + t^2, \qquad -L < t < L$

9. $f(t) = \begin{cases} 0, & -\pi \le t \le 0 \\ \sin(t), & 0 \le t \le \pi \end{cases}$

10. $f(t) = \begin{cases} t, & -\frac{1}{2} \le t \le \frac{1}{2} \\ 1 - t, & \frac{1}{2} \le t \le \frac{3}{2} \end{cases}$

11. $f(t) = \begin{cases} 0, & -a < t < 0 \\ 2t, & 0 < t < a \end{cases}$

12. $f(t) = \begin{cases} 0, & -\pi < t \le 0 \\ t^2, & 0 \le t < \pi \end{cases}$

13. $f(t) = (\pi - t)/2, \quad 0 < t < 2$

14. $f(t) = t\cos(\pi t/L)\,, \; -L < t < L$

15. $f(t) = \sinh[a\,(\pi/2 - |t|)]\,, \; -\pi \le t \le \pi$

16. $f(t) = \begin{cases} x(2L - x), & 0 \le t \le 2L \\ x^2 - 6Lx + 8L^2, & 2L \le t \le 4L \end{cases}$

5.2 PROPERTIES OF FOURIER SERIES

In the previous section we introduced the Fourier series and showed how to compute one given the function $f(t)$. In this section we examine some particular properties of these series.

---

[8] White noise is sound that is analogous to white light in that it is uniformly distributed throughout the complete audible sound spectrum.

Differentiation of a Fourier series

In certain instances we only have the Fourier series representation of a function $f(t)$. Can we find the derivative or the integral of $f(t)$ merely by differentiating or integrating the Fourier series term by term? Is this permitted? Let us consider the case of differentiation first.

Consider a function $f(t)$ of period $2L$, which has the derivative $f'(t)$. Let us assume that we can expand $f'(t)$ as a Fourier series. This implies that $f'(t)$ is continuous except for a finite number of discontinuities and $f(t)$ is continuous over an interval that starts at $t = \tau$ and ends at $t = \tau + 2L$. Then

$$f'(t) = \frac{a_0'}{2} + \sum_{n=1}^{\infty} a_n' \cos\left(\frac{n\pi t}{L}\right) + b_n' \sin\left(\frac{n\pi t}{L}\right), \tag{5.2.1}$$

where we denoted the Fourier coefficients of $f'(t)$ with a prime. Computing the Fourier coefficients,

$$a_0' = \frac{1}{L}\int_{\tau}^{\tau+2L} f'(t)\,dt = \frac{1}{L}\left[f(\tau+2L) - f(\tau)\right] = 0, \tag{5.2.2}$$

if $f(\tau + 2L) = f(\tau)$. Similarly, by integrating by parts,

$$a_n' = \frac{1}{L}\int_{\tau}^{\tau+2L} f'(t)\cos\left(\frac{n\pi t}{L}\right)\,dt \tag{5.2.3}$$

$$= \frac{1}{L}\left[f(t)\cos\left(\frac{n\pi t}{L}\right)\right]\Bigg|_{\tau}^{\tau+2L} + \frac{n\pi}{L^2}\int_{\tau}^{\tau+2L} f(t)\sin\left(\frac{n\pi t}{L}\right)\,dt \tag{5.2.4}$$

$$= \frac{n\pi b_n}{L}, \tag{5.2.5}$$

and

$$b_n' = \frac{1}{L}\int_{\tau}^{\tau+2L} f'(t)\sin\left(\frac{n\pi t}{L}\right)\,dt \tag{5.2.6}$$

$$= \frac{1}{L}\left[f(t)\sin\left(\frac{n\pi t}{L}\right)\right]\Bigg|_{\tau}^{\tau+2L} - \frac{n\pi}{L^2}\int_{\tau}^{\tau+2L} f(t)\cos\left(\frac{n\pi t}{L}\right)\,dt \tag{5.2.7}$$

$$= -\frac{n\pi a_n}{L}. \tag{5.2.8}$$

Consequently, if we have a function $f(t)$ whose derivative $f'(t)$ is continuous except for a finite number of discontinuities and $f(\tau) = f(\tau + 2L)$, then

$$f'(t) = \sum_{n=1}^{\infty} \frac{n\pi}{L}\left[b_n \cos\left(\frac{n\pi t}{L}\right) - a_n \sin\left(\frac{n\pi t}{L}\right)\right]. \tag{5.2.9}$$

That is, the derivative of $f(t)$ is given by a term-by-term differentiation of the Fourier series of $f(t)$.

• Example 5.2.1

The Fourier series for the periodic function

$$f(t) = \begin{cases} 0, & -\pi \le t \le 0, \\ t, & 0 \le t \le \pi/2, \\ \pi - t, & \pi/2 \le t \le \pi, \end{cases} \qquad f(t) = f(t + 2\pi), \tag{5.2.10}$$

is

$$f(t) = \frac{\pi}{8} - \frac{1}{\pi}\sum_{n=1}^{\infty} \frac{\cos[2(2n-1)t]}{(2n-1)^2} - \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)^2}\sin[(2n-1)t]. \tag{5.2.11}$$

Because $f(t)$ is continuous over the entire interval $(-\pi, \pi)$ and $f(-\pi) = f(\pi) = 0$, we can find $f'(t)$ by taking the derivative of Equation 5.2.11 term by term:

$$f'(t) = \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{\sin[2(2n-1)t]}{2n-1} - \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{(-1)^n}{2n-1}\cos[(2n-1)t]. \tag{5.2.12}$$

This is the same Fourier series that we would obtain by computing the Fourier series for

$$f'(t) = \begin{cases} 0, & -\pi < t < 0, \\ 1, & 0 < t < \pi/2, \\ -1, & \pi/2 < t < \pi. \end{cases} \tag{5.2.13}$$

□

Integration of a Fourier series

To determine whether we can find the integral of $f(t)$ by term-by-term integration of its Fourier series, consider a form of the antiderivative of $f(t)$:

$$F(t) = \int_0^t \left[f(\tau) - \frac{a_0}{2} \right] d\tau. \tag{5.2.14}$$

Now

$$F(t+2L) = \int_0^t \left[f(\tau) - \frac{a_0}{2} \right] d\tau + \int_t^{t+2L} \left[f(\tau) - \frac{a_0}{2} \right] d\tau \tag{5.2.15}$$

$$= F(t) + \int_{-L}^{L} \left[f(\tau) - \frac{a_0}{2} \right] d\tau \tag{5.2.16}$$

$$= F(t) + \int_{-L}^{L} f(\tau)\, d\tau - La_0 = F(t), \tag{5.2.17}$$

so that $F(t)$ has a period of $2L$. Consequently we may expand $F(t)$ as the Fourier series

$$F(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi t}{L}\right) + B_n \sin\left(\frac{n\pi t}{L}\right). \tag{5.2.18}$$

For A_n,

$$A_n = \frac{1}{L}\int_{-L}^{L} F(t)\cos\left(\frac{n\pi t}{L}\right)\,dt \tag{5.2.19}$$

$$= \frac{1}{L}\left[F(t)\frac{\sin(n\pi t/L)}{n\pi/L}\right]\Bigg|_{-L}^{L} - \frac{1}{n\pi}\int_{-L}^{L}\left[f(t) - \frac{a_0}{2}\right]\sin\left(\frac{n\pi t}{L}\right)\,dt \tag{5.2.20}$$

$$= -\frac{b_n}{n\pi/L}. \tag{5.2.21}$$

Similarly,

$$B_n = \frac{a_n}{n\pi/L}. \tag{5.2.22}$$

Therefore,

$$\int_0^t f(\tau)\,d\tau = \frac{a_0 t}{2} + \frac{A_0}{2} + \sum_{n=1}^{\infty}\frac{a_n\sin(n\pi t/L) - b_n\cos(n\pi t/L)}{n\pi/L}. \tag{5.2.23}$$

This is identical to a term-by-term integration of the Fourier series for $f(t)$. Thus, we can always find the integral of $f(t)$ by a term-by-term integration of its Fourier series.

• **Example 5.2.2**

The Fourier series for $f(t) = t$ for $-\pi < t < \pi$ is

$$f(t) = -2\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin(nt). \tag{5.2.24}$$

To find the Fourier series for $f(t) = t^2$, we integrate Equation 5.2.24 term by term and find that

$$\frac{\tau^2}{2}\Bigg|_0^t = 2\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nt) - 2\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}. \tag{5.2.25}$$

But $\sum_{n=1}^{\infty}(-1)^n/n^2 = -\pi^2/12$. Substituting and multiplying by 2, we obtain the final result that

$$t^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nt). \tag{5.2.26}$$

□

Parseval's equality

One of the fundamental quantities in engineering is power. The *power content* of a periodic signal $f(t)$ of period $2L$ is $\int_{\tau}^{\tau+2L} f^2(t)\,dt/L$. This mathematical definition mirrors the power dissipation I^2R that occurs in a resistor of resistance R where I is the root mean square (RMS) of the current. We would like to compute this power content as simply as possible given the coefficients of its Fourier series.

Assume that $f(t)$ has the Fourier series

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right). \tag{5.2.27}$$

Then,

$$\frac{1}{L}\int_{\tau}^{\tau+2L} f^2(t)\,dt = \frac{a_0}{2L}\int_{\tau}^{\tau+2L} f(t)\,dt + \sum_{n=1}^{\infty}\frac{a_n}{L}\int_{\tau}^{\tau+2L} f(t)\cos\left(\frac{n\pi t}{L}\right)\,dt$$

$$+ \sum_{n=1}^{\infty}\frac{b_n}{L}\int_{\tau}^{\tau+2L} f(t)\sin\left(\frac{n\pi t}{L}\right)\,dt \tag{5.2.28}$$

$$= \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2). \tag{5.2.29}$$

Equation 5.2.29 is *Parseval's equality.*[9] It allows us to sum squares of Fourier coefficients (which we have already computed) rather than performing the integration $\int_{\tau}^{\tau+2L} f^2(t)\,dt$ analytically or numerically.

• Example 5.2.3

The Fourier series for $f(t) = t^2$ over the interval $[-\pi, \pi]$ is

$$t^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nt). \tag{5.2.30}$$

Then, by Parseval's equality,

$$\frac{1}{\pi}\int_{-\pi}^{\pi} t^4\,dt = \left.\frac{2t^5}{5\pi}\right|_0^{\pi} = \frac{4\pi^4}{18} + 16\sum_{n=1}^{\infty}\frac{1}{n^4}, \quad \text{or} \quad \left(\frac{2}{5} - \frac{4}{18}\right)\pi^4 = 16\sum_{n=1}^{\infty}\frac{1}{n^4}. \tag{5.2.31}$$

Consequently,

$$\pi^4/90 = \sum_{n=1}^{\infty}\frac{1}{n^4}. \tag{5.2.32}$$

□

Gibbs phenomena

In the actual application of Fourier series, we cannot sum an infinite number of terms but must be content with N terms. If we denote this partial sum of the Fourier series by

---

[9] Parseval, M.-A., 1805: Mémoire sur les séries et sur l'intégration complète d'une équation aux différences partielles linéaires du second ordre, à coefficients constants. *Mémoires présentés a l'Institut des sciences, lettres et arts, par divers savans, et lus dans ses assemblées: Sciences mathématiques et Physiques*, **1**, 638–648.

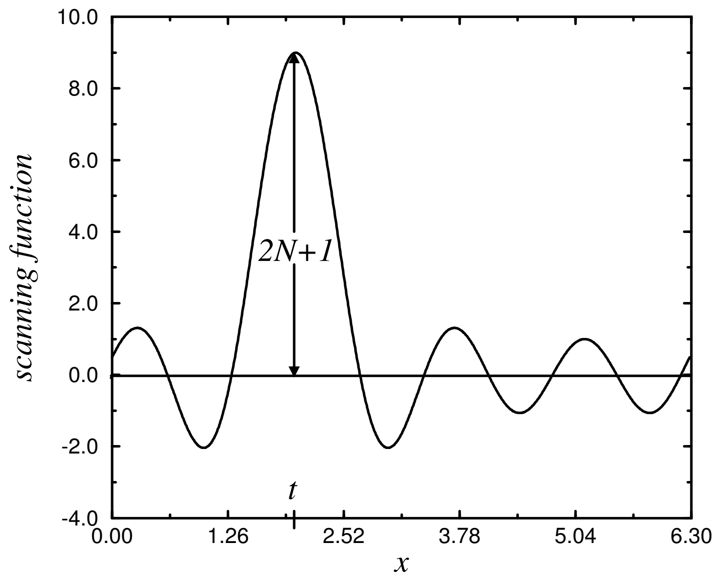


Figure 5.2.1: The scanning function over $0 \le x \le 2\pi$ for $N = 5$.

$S_N(t)$, we have from the definition of the Fourier series:

$$S_N(t) = \tfrac{1}{2}a_0 + \sum_{n=1}^{N} a_n \cos(nt) + b_n \sin(nt) \qquad (\mathbf{5.2.33})$$

$$= \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx + \frac{1}{\pi}\int_0^{2\pi} f(x)\left[\sum_{n=1}^{N} \cos(nt)\cos(nx) + \sin(nt)\sin(nx)\right] dx \qquad (\mathbf{5.2.34})$$

$$= \frac{1}{\pi}\int_0^{2\pi} f(x)\left\{\frac{1}{2} + \sum_{n=1}^{N} \cos[n(t-x)]\right\} dx \qquad (\mathbf{5.2.35})$$

$$= \frac{1}{2\pi}\int_0^{2\pi} f(x)\frac{\sin[(N+\frac{1}{2})(x-t)]}{\sin[\frac{1}{2}(x-t)]}\,dx. \qquad (\mathbf{5.2.36})$$

The quantity $\sin[(N+\frac{1}{2})(x-t)]/\sin[\frac{1}{2}(x-t)]$ is called a *scanning function.* Over the range $0 \le x \le 2\pi$ it has a very large peak at $x = t$ where the amplitude equals $2N+1$. See Figure 5.2.1. On either side of this peak there are oscillations that decrease rapidly with distance from the peak. Consequently, as $N \to \infty$, the scanning function becomes essentially a long narrow slit corresponding to the area under the large peak at $x = t$. If we neglect for the moment the small area under the minor ripples adjacent to this slit, then the integral, Equation 5.2.36, essentially equals $f(t)$ times the area of the slit divided by 2π. If $1/2\pi$ times the area of the slit equals unity, then the value of $S_N(t) \approx f(t)$ to a good approximation for large N.

For relatively small values of N, the scanning function deviates considerably from its ideal form, and the partial sum $S_N(t)$ only crudely approximates $f(t)$. As the partial sum includes more terms and N becomes relatively large, the form of the scanning function improves and so does the agreement between $S_N(t)$ and $f(t)$. The improvement in the scanning function is due to the large hump becoming taller and narrower. At the same time, the adjacent ripples become more numerous as well as narrower in the same proportion as the large hump does.

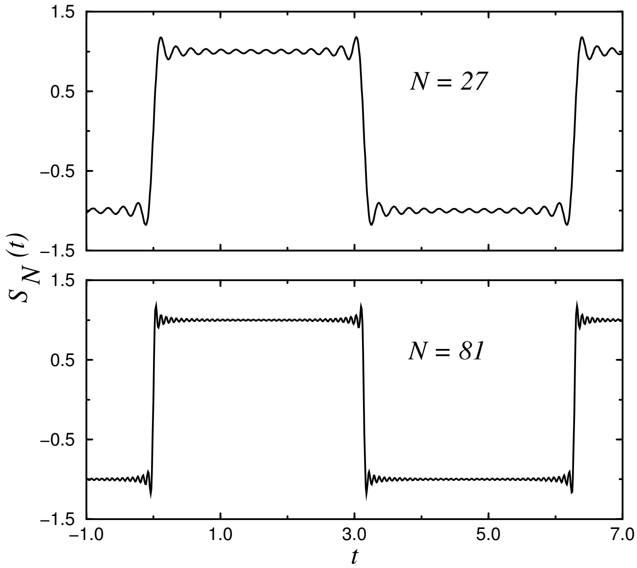


Figure 5.2.2: The finite Fourier series representation $S_N(t)$ for the function, Equation 5.2.38, for the range $-1 \le t \le 7$ for $N = 27$ and $N = 81$.

The reason why $S_N(t)$ and $f(t)$ will never become identical, even in the limit of $N \to \infty$, is the presence of the positive and negative side lobes near the large peak. Because

$$\frac{\sin[(N+\frac{1}{2})(x-t)]}{\sin[\frac{1}{2}(x-t)]} = 1 + 2\sum_{n=1}^{N} \cos[n(t-x)], \tag{5.2.37}$$

an integration of the scanning function over the interval 0 to 2π shows that the total area under the scanning function equals 2π. However, from Figure 5.2.1 the net area contributed by the ripples is numerically negative so that the area under the large peak must exceed 2π if the total area equals 2π. Although the exact value depends upon N, it is important to note that this excess does not become zero as $N \to \infty$.

Thus, the presence of these negative side lobes explains the departure of our scanning function from the idealized slit of area 2π. To illustrate this departure, consider the function:

$$f(t) = \begin{cases} 1, & 0 < t < \pi, \\ -1, & \pi < t < 2\pi. \end{cases} \tag{5.2.38}$$

Then,

$$S_N(t) = \frac{1}{2\pi}\int_0^{\pi} \frac{\sin[(N+\frac{1}{2})(x-t)]}{\sin[\frac{1}{2}(x-t)]}\,dx - \frac{1}{2\pi}\int_{\pi}^{2\pi} \frac{\sin[(N+\frac{1}{2})(x-t)]}{\sin[\frac{1}{2}(x-t)]}\,dx \tag{5.2.39}$$

$$= \frac{1}{2\pi}\int_0^{\pi} \left\{ \frac{\sin[(N+\frac{1}{2})(x-t)]}{\sin[\frac{1}{2}(x-t)]}\,dx + \frac{\sin[(N+\frac{1}{2})(x+t)]}{\sin[\frac{1}{2}(x+t)]}\,dx \right\} \tag{5.2.40}$$

$$= \frac{1}{2\pi}\int_{-t}^{\pi-t} \frac{\sin[(N+\frac{1}{2})\theta]}{\sin(\frac{1}{2}\theta)}\,d\theta - \frac{1}{2\pi}\int_{t}^{\pi+t} \frac{\sin[(N+\frac{1}{2})\theta]}{\sin(\frac{1}{2}\theta)}\,d\theta. \tag{5.2.41}$$

The first integral in Equation 5.2.41 gives the contribution to $S_N(t)$ from the jump discontinuity at $t = 0$ while the second integral gives the contribution from $t = \pi$. In Figure

5.2.2 we have plotted $S_N(t)$ when $N = 27$ and $N = 81$. Residual discrepancies remain even for very large values of N. Indeed, as N increases, this figure changes only in that the ripples in the vicinity of the discontinuity of $f(t)$ proportionally increase their rate of oscillation as a function of t while their relative magnitude remains the same. As $N \to \infty$ these ripples compress into a single vertical line at the point of discontinuity. True, these oscillations occupy smaller and smaller spaces but they still remain. Thus, we can never approximate a function in the vicinity of a discontinuity by a finite Fourier series without suffering from this over- and undershooting of the series. This peculiarity of Fourier series is called the *Gibbs phenomena*.[10] Gibbs phenomena can only be eliminated by removing the discontinuity.[11]

Problems

Additional Fourier series representations can be generated by differentiating or integrating known Fourier series. Work out the following two examples.

1. Given

$$\frac{\pi^2 - 2\pi x}{8} = \sum_{n=0}^{\infty} \frac{\cos[(2n+1)x]}{(2n+1)^2}, \qquad 0 \le x \le \pi,$$

obtain

$$\frac{\pi^2 x - \pi x^2}{8} = \sum_{n=0}^{\infty} \frac{\sin[(2n+1)x]}{(2n+1)^3}, \qquad 0 \le x \le \pi,$$

by term-by-term integration. Could we go the other way, i.e., take the derivative of the second equation to obtain the first? Explain.

2. Given

$$\frac{\pi^2 - 3x^2}{12} = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\cos(nx)}{n^2}, \qquad -\pi \le x \le \pi,$$

obtain

$$\frac{\pi^2 x - x^3}{12} = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin(nx)}{n^3}, \qquad -\pi \le x \le \pi,$$

by term-by-term integration. Could we go the other way, i.e., take the derivative of the second equation to obtain the first? Explain.

3. (a) Show that the Fourier series for the odd function:

$$f(t) = \begin{cases} 2t + t^2, & -2 \le t \le 0, \\ 2t - t^2, & 0 \le t \le 2, \end{cases} \quad \text{is} \quad f(t) = \frac{32}{\pi^3} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^3} \sin\left[\frac{(2n-1)\pi t}{2}\right].$$

[10] Gibbs, J. W., 1898: Fourier's series. *Nature*, **59**, 200; Gibbs, J. W., 1899: Fourier's series. *Nature*, **59**, 606. For the historical development, see Hewitt, E., and R. E. Hewitt, 1979: The Gibbs-Wilbraham phenomenon: An episode in Fourier analysis. *Arch. Hist. Exact Sci.*, **21**, 129–160.

[11] For a particularly clever method for improving the convergence of a trigonometric series, see Kantorovich, L. V., and V. I. Krylov, 1964: *Approximate Methods of Higher Analysis*. Interscience, pp. 77–88.

(b) Use Parseval's equality to show that

$$\frac{\pi^6}{960} = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^6}.$$

This series converges very rapidly to $\pi^6/960$ and provides a convenient method for computing π^6.

5.3 HALF-RANGE EXPANSIONS

In certain applications, we will find that we need a Fourier series representation for a function $f(x)$ that applies over the interval $(0, L)$ rather than $(-L, L)$. Because we are completely free to define the function over the interval $(-L, 0)$, it is simplest to have a series that consists only of sines or cosines. In this section we shall show how we can obtain these so-called *half-range expansions*.

Recall in Example 5.1.3 how we saw that if $f(x)$ is an even function, then $b_n = 0$ for all n. Similarly, if $f(x)$ is an odd function, then $a_0 = a_n = 0$ for all n. We now use these results to find a Fourier half-range expansion by extending the function defined over the interval $(0, L)$ as either an even or odd function into the interval $(-L, 0)$. If we extend $f(x)$ as an even function, we will get a half-range cosine series; if we extend $f(x)$ as an odd function, we obtain a half-range sine series.

It is important to remember that half-range expansions are a special case of the general Fourier series. For any $f(x)$ we can construct either a Fourier sine or cosine series over the interval $(-L, L)$. Both of these series will give the correct answer over the interval of $(0, L)$. Which one we choose to use depends upon whether we wish to deal with a cosine or sine series.

• Example 5.3.1

Let us find the half-range sine expansion of

$$f(x) = 1, \qquad 0 < x < \pi. \tag{5.3.1}$$

We begin by defining the periodic odd function

$$\widetilde{f}(x) = \begin{cases} -1, & -\pi < x < 0, \\ 1, & 0 < x < \pi, \end{cases} \tag{5.3.2}$$

with $\widetilde{f}(x + 2\pi) = \widetilde{f}(x)$. Because $\widetilde{f}(x)$ is odd, $a_0 = a_n = 0$ and

$$b_n = \frac{2}{\pi} \int_0^{\pi} 1 \sin(nx)\, dx = -\frac{2}{n\pi} \cos(nx)\Big|_0^{\pi} = -\frac{2}{n\pi}\left[\cos(n\pi) - 1\right] = -\frac{2}{n\pi}\left[(-1)^n - 1\right]. \tag{5.3.3}$$

The Fourier half-range sine series expansion of $f(x)$ is therefore

$$f(x) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{[1 - (-1)^n]}{n} \sin(nx) = \frac{4}{\pi} \sum_{m=1}^{\infty} \frac{\sin[(2m-1)x]}{2m-1}. \tag{5.3.4}$$

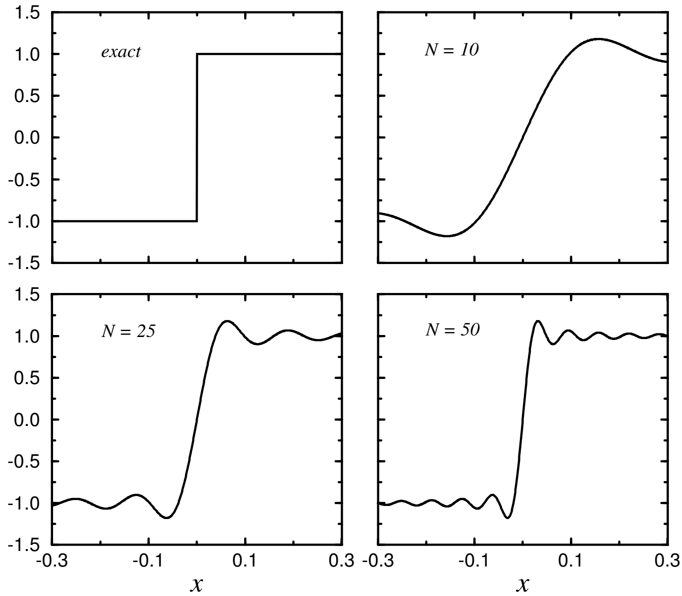


Figure 5.3.1: Partial sum of N terms in the Fourier half-range sine representation of a square wave.

As counterpoint, let us find the half-range cosine expansion of $f(x) = 1$, $0 < x < \pi$. Now, we have that $b_n = 0$,

$$a_0 = \frac{2}{\pi}\int_0^\pi 1\,dx = 2, \tag{5.3.5}$$

and

$$a_n = \frac{2}{\pi}\int_0^\pi \cos(nx)\,dx = \frac{2}{n\pi}\sin(nx)\Big|_0^\pi = 0. \tag{5.3.6}$$

Thus, the Fourier half-range cosine expansion equals the single term:

$$f(x) = 1, \qquad 0 < x < \pi. \tag{5.3.7}$$

This is perfectly reasonable. To form a half-range cosine expansion we extend $f(x)$ as an even function into the interval $(-\pi, 0)$. In this case, we would obtain $\widetilde{f}(x) = 1$ for $-\pi < x < \pi$. Finally, we note that the Fourier series of a constant is simply that constant.

In practice it is impossible to sum Equation 5.3.4 exactly and we actually sum only the first N terms. Figure 5.3.1 illustrates $f(x)$ when this Fourier series contains N terms. As seen from the figure, the truncated series tries to achieve the infinite slope at $x = 0$, but in the attempt, it *overshoots* the discontinuity by a certain amount (in this particular case, by 17.9%). This is another example of the Gibbs phenomena. Increasing the number of terms does not remove this peculiarity; it merely shifts it nearer to the discontinuity. □

• Example 5.3.2: Inertial supercharging of an engine

An important aspect of designing any gasoline engine involves the motion of the fuel, air, and exhaust gas mixture through the engine. Ordinarily an engineer would consider the motion as steady flow; but in the case of a four-stroke, single-cylinder gasoline engine, the closing of the intake valve interrupts the steady flow of the gasoline-air mixture for nearly three quarters of the engine cycle. This periodic interruption sets up standing waves in the intake pipe - waves that can build up an appreciable pressure amplitude just outside the input value.

When one of the harmonics of the engine frequency equals one of the resonance frequencies of the intake pipe, then the pressure fluctuations at the valve will be large. If the intake valve closes during that portion of the cycle when the pressure is less than average, then the waves will reduce the power output. However, if the intake valve closes when the pressure is greater than atmospheric, then the waves will have a supercharging effect and will produce an increase of power. This effect is called *inertia supercharging*.

While studying this problem, Morse et al.[12] found it necessary to express the velocity of the air-gas mixture in the valve, given by

$$f(t) = \begin{cases} 0, & -\pi < \omega t < -\pi/4, \\ \pi\cos(2\omega t)/2, & -\pi/4 < \omega t < \pi/4, \\ 0, & \pi/4 < \omega t < \pi, \end{cases} \tag{5.3.8}$$

in terms of a Fourier expansion. The advantage of working with the Fourier series rather than the function itself lies in the ability to write the velocity as a periodic forcing function that highlights the various harmonics that might resonate with the structure comprising the fuel line.

Clearly $f(t)$ is an even function and its Fourier representation will be a cosine series. In this problem $\tau = -\pi/\omega$, and $L = \pi/\omega$. Therefore,

$$a_0 = \frac{2\omega}{\pi}\int_{-\pi/4\omega}^{\pi/4\omega} \frac{\pi}{2}\cos(2\omega t)\,dt = \tfrac{1}{2}\sin(2\omega t)\Big|_{-\pi/4\omega}^{\pi/4\omega} = 1, \tag{5.3.9}$$

and

$$a_n = \frac{2\omega}{\pi}\int_{-\pi/4\omega}^{\pi/4\omega} \frac{\pi}{2}\cos(2\omega t)\cos\left(\frac{n\pi t}{\pi/\omega}\right)dt \tag{5.3.10}$$

$$= \frac{\omega}{2}\int_{-\pi/4\omega}^{\pi/4\omega} \{\cos[(n+2)\omega t] + \cos[(n-2)\omega t]\}\,dt \tag{5.3.11}$$

$$= \begin{cases} \left.\frac{\sin[(n+2)\omega t]}{2(n+2)} + \frac{\sin[(n-2)\omega t]}{2(n-2)}\right|_{-\pi/4\omega}^{\pi/4\omega}, & n \neq 2, \\ \left.\frac{\omega t}{2} + \frac{\sin(4\omega t)}{4}\right|_{-\pi/4\omega}^{\pi/4\omega}, & n = 2, \end{cases} \tag{5.3.12}$$

$$= \begin{cases} -\frac{4}{n^2-4}\cos\left(\frac{n\pi}{4}\right), & n \neq 2, \\ \frac{\pi}{4}, & n = 2. \end{cases} \tag{5.3.13}$$

Plotting these Fourier coefficients using the MATLAB script:

```
for m = 1:21;
  n = m-1; % compute the indices for the harmonic
% compute the Fourier coefficients a_n
  if n == 2; an(m) = pi/4; else;
  an(m) = 4.*cos(pi*n/4)/(4-n*n); end;
```

[12] Morse, P. M., R. H. Boden, and H. Schecter, 1938: Acoustic vibrations and internal combustion engine performance. I. Standing waves in the intake pipe system. *J. Appl. Phys.*, **9**, 16–23.

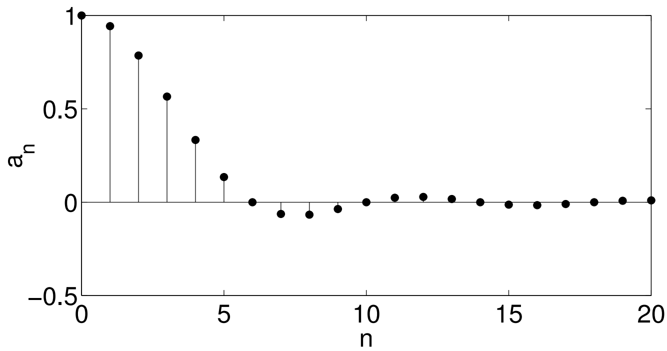


Figure 5.3.2: The spectral coefficients of the Fourier cosine series of the function given by Equation 5.3.9.

```
end
nn=0:20; % create indices for x-axis
fzero=zeros(size(nn)); % create the zero line
clf % clear any figures
axes('FontSize',20) % set font size
stem(nn,an,'filled') % plot spectrum
hold on
plot(nn,fzero,'-') % plot the zero line
set (gca,'PlotBoxAspectRatio',[8 4 1]) % set aspect ratio
xlabel('n') % label x-axis
ylabel('a_n') % label y-axis,
```

we see that these Fourier coefficients become small rapidly (see Figure 5.3.2). For that reason, Morse et al. showed that there are only about three resonances where the acoustic properties of the intake pipe can enhance engine performance. These peaks occur when $q = 30c/NL = 3, 4$, or 5, where c is the velocity of sound in the air-gas mixture, L is the effective length of the intake pipe, and N is the engine speed in rpm. See Figure 5.3.3. Subsequent experiments[13] verified these results.

Such analyses are valuable to automotive engineers. Engineers are always seeking ways to optimize a system with little or no additional cost. Our analysis shows that by tuning the length of the intake pipe so that it falls on one of the resonance peaks, we could obtain higher performance from the engine with little or no extra work. Of course, the problem is that no car always performs at some optimal condition.

Problems

Find the Fourier cosine and sine series for the following functions. Then, use MATLAB to plot the Fourier coefficients.

1. $f(t) = t, \qquad 0 < t < \pi$

2. $f(t) = \pi - t, \qquad 0 < t < \pi$

3. $f(t) = t(a - t), \qquad 0 < t < a$

4. $f(t) = e^{kt}, \qquad 0 < t < a$

---

[13] Boden, R. H., and H. Schecter, 1944: Dynamics of the inlet system of a four-stroke engine. *NACA Tech. Note 935*.

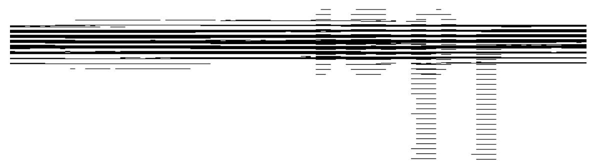


Figure 5.3.3: Experimental verification of inertial supercharging within a gasoline engine resulting from the resonance of the air-gas mixture and the intake pipe system. The peaks correspond to the $n = 3, 4$, and 5 harmonics of the Fourier representation, Equation 5.3.13, and the parameter q is defined in the text. (From Morse, P., R. H. Boden, and H. Schecter, 1938: Acoustic vibrations and internal combustion engine performance. *J. Appl. Phys.*, **9**, 17 with permission.)

5. $f(t) = \begin{cases} t, & 0 \le t \le \frac{1}{2} \\ 1-t, & \frac{1}{2} \le t \le 1 \end{cases}$

6. $f(t) = \begin{cases} t, & 0 < t \le 1 \\ 1, & 1 \le t < 2 \end{cases}$

7. $f(t) = \pi^2 - t^2, \quad 0 < t < \pi$

8. $f(t) = \begin{cases} 0, & 0 < t < \dfrac{a}{2} \\ 1, & \dfrac{a}{2} < t < a \end{cases}$

9. $f(t) = \begin{cases} 0, & 0 < t \le \dfrac{a}{3} \\ t - \dfrac{a}{3}, & \dfrac{a}{3} \le t \le \dfrac{2a}{3} \\ \dfrac{a}{3}, & \dfrac{2a}{3} \le t < a \end{cases}$

10. $f(t) = \begin{cases} 0, & 0 < t < \dfrac{a}{4} \\ 1, & \dfrac{a}{4} < t < \dfrac{3a}{4} \\ 0, & \dfrac{3a}{4} < t < a \end{cases}$

11. $f(t) = \begin{cases} \dfrac{1}{2}, & 0 < t < \dfrac{a}{2} \\ 1, & \dfrac{a}{2} < t < a \end{cases}$

12. $f(t) = \begin{cases} \dfrac{2t}{a}, & 0 < t \le \dfrac{a}{2} \\ \dfrac{3a-2t}{2a}, & \dfrac{a}{2} \le t < a \end{cases}$

13. $f(t) = \begin{cases} t, & 0 < t \le \dfrac{a}{2} \\ \dfrac{a}{2}, & \dfrac{a}{2} \le t < a \end{cases}$

14. $f(t) = \dfrac{a-t}{a}, \quad 0 < t < a$

15. Using the relationships[14] that

[14] Gradshteyn, I. S., and I. M. Ryzhik, 1965: *Table of Integrals, Series, and Products*. Academic Press, Section 3.753, Formula 2 and Section 3.771, Formula 8.

$$\int_0^1 \frac{\cos(ax)}{\sqrt{1-x^2}}\,dx = \frac{\pi}{2}J_0(a), \text{ and } \int_0^u (u^2-x^2)^{\nu-\frac{1}{2}}\cos(ax)\,dx = \frac{\sqrt{\pi}}{2}\left(\frac{2u}{a}\right)^{\nu}\Gamma\left(\nu+\tfrac{1}{2}\right)J_\nu(au),$$

with $a > 0$, $u > 0$, $\Re(\nu) > -\frac{1}{2}$, obtain the following half-range expansions:

$$\frac{1}{\sqrt{1-x^2}} = \frac{\pi}{2} + \pi\sum_{n=1}^{\infty} J_0(n\pi)\cos(n\pi x), \qquad 0 < x < 1,$$

and

$$\sqrt{1-x^2} = 2\sum_{n=1}^{\infty}\frac{J_1[(2n-1)\pi/2]}{2n-1}\cos[(2n-1)\pi x/2], \qquad 0 < x < 1.$$

Here $J_\nu(\cdot)$ denotes the Bessel function of the first kind and order ν (see Section 6.5) and $\Gamma(\cdot)$ is the gamma function.[15]

16. The function

$$f(t) = 1 - (1+a)\frac{t}{\pi} + (a-1)\frac{t^2}{\pi^2} + (a+1)\frac{t^3}{\pi^3} - a\frac{t^4}{\pi^4}, \qquad 0 < t < \pi$$

is a curve fit to the observed pressure trace of an explosion wave in the atmosphere. Because the observed transmission of atmospheric waves depends on the five-fourths power of the frequency, Reed[16] had to re-express this curve fit as a Fourier sine series before he could use the transmission law. He found that

$$\begin{aligned} f(t) &= \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\left[1 - \frac{3(a-1)}{2\pi^2 n^2}\right]\sin(2nt) \\ &+ \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{2}{2n-1}\left[1 + \frac{2(a-1)}{\pi^2(2n-1)^2} - \frac{48a}{\pi^4(2n-1)^4}\right]\sin[(2n-1)t]. \end{aligned}$$

Confirm his result.

5.4 FOURIER SERIES WITH PHASE ANGLES

Sometimes it is desirable to rewrite a general Fourier series as a purely cosine or purely sine series with a phase angle. Engineers often speak of some quantity leading or lagging another quantity. Re-expressing a Fourier series in terms of amplitude and phase provides a convenient method for determining these phase relationships.

Suppose, for example, that we have a function $f(t)$ of period $2L$, given in the interval $[-L, L]$, whose Fourier series expansion is

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos\left(\frac{n\pi t}{L}\right) + b_n\sin\left(\frac{n\pi t}{L}\right). \qquad \textbf{(5.4.1)}$$

---

[15] Gradshteyn and Ryzhik, op. cit., Section 6.41.

[16] Reed, J. W., 1977: Atmospheric attenuation of explosion waves. *J. Acoust. Soc. Am.*, **61**, 39–47.

We wish to replace Equation 5.4.1 by the series:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi t}{L} + \varphi_n\right). \tag{5.4.2}$$

To do this we note that

$$B_n \sin\left(\frac{n\pi t}{L} + \varphi_n\right) = a_n \cos\left(\frac{n\pi t}{L}\right) + b_n \sin\left(\frac{n\pi t}{L}\right) \tag{5.4.3}$$

$$= B_n \sin\left(\frac{n\pi t}{L}\right)\cos(\varphi_n) + B_n \sin(\varphi_n)\cos\left(\frac{n\pi t}{L}\right). \tag{5.4.4}$$

We equate coefficients of $\sin(n\pi t/L)$ and $\cos(n\pi t/L)$ on both sides and obtain

$$a_n = B_n \sin(\varphi_n), \qquad \text{and} \qquad b_n = B_n \cos(\varphi_n). \tag{5.4.5}$$

Hence, upon squaring and adding,

$$B_n = \sqrt{a_n^2 + b_n^2}, \tag{5.4.6}$$

while taking the ratio gives

$$\varphi_n = \tan^{-1}(a_n/b_n). \tag{5.4.7}$$

Similarly we could rewrite Equation 5.4.1 as

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi t}{L} + \varphi_n\right), \tag{5.4.8}$$

where

$$A_n = \sqrt{a_n^2 + b_n^2}, \qquad \text{and} \qquad \varphi_n = \tan^{-1}(-b_n/a_n), \tag{5.4.9}$$

and

$$a_n = A_n \cos(\varphi_n), \qquad \text{and} \qquad b_n = -A_n \sin(\varphi_n). \tag{5.4.10}$$

In both cases, we must be careful in computing φ_n because there are two possible values of φ_n that satisfy Equation 5.4.7 or Equation 5.4.9. These angles φ_n must give the correct a_n and b_n using either Equation 5.4.5 or Equation 5.4.10.

• Example 5.4.1

The Fourier series for $f(t) = e^t$ over the interval $-L < t < L$ is

$$f(t) = \frac{\sinh(aL)}{aL} + 2\sinh(aL)\sum_{n=1}^{\infty} \frac{aL(-1)^n}{a^2L^2 + n^2\pi^2}\cos\left(\frac{n\pi t}{L}\right)$$
$$- 2\sinh(aL)\sum_{n=1}^{\infty} \frac{n\pi(-1)^n}{a^2L^2 + n^2\pi^2}\sin\left(\frac{n\pi t}{L}\right). \tag{5.4.11}$$

Let us rewrite Equation 5.4.11 as a Fourier series with a phase angle. Regardless of whether we want the new series to contain $\cos(n\pi t/L + \varphi_n)$ or $\sin(n\pi t/L + \varphi_n)$, the amplitude A_n or B_n is the same in both series:

$$A_n = B_n = \sqrt{a_n^2 + b_n^2} = \frac{2\sinh(aL)}{\sqrt{a^2L^2 + n^2\pi^2}}. \tag{5.4.12}$$

If we want our Fourier series to read

$$f(t) = \frac{\sinh(aL)}{aL} + 2\sinh(aL)\sum_{n=1}^{\infty}\frac{\cos(n\pi t/L + \varphi_n)}{\sqrt{a^2L^2 + n^2\pi^2}}, \qquad \textbf{(5.4.13)}$$

then

$$\varphi_n = \tan^{-1}\left(-\frac{b_n}{a_n}\right) = \tan^{-1}\left(\frac{n\pi}{aL}\right), \qquad \textbf{(5.4.14)}$$

where φ_n lies in the first quadrant if n is even and in the third quadrant if n is odd. This ensures that the sign from the $(-1)^n$ is correct.

On the other hand, if we prefer

$$f(t) = \frac{\sinh(aL)}{aL} + 2\sinh(aL)\sum_{n=1}^{\infty}\frac{\sin(n\pi t/L + \varphi_n)}{\sqrt{a^2L^2 + n^2\pi^2}}, \qquad \textbf{(5.4.15)}$$

then

$$\varphi_n = \tan^{-1}\left(\frac{a_n}{b_n}\right) = -\tan^{-1}\left(\frac{aL}{n\pi}\right), \qquad \textbf{(5.4.16)}$$

where φ_n lies in the fourth quadrant if n is odd and in the second quadrant if n is even.

Problems

Write the following Fourier series in both the cosine and sine phase angle form:

1. $f(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(2n-1)\pi t]}{2n-1}$

2. $f(t) = \frac{3}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{2n-1}\cos\left[\frac{(2n-1)\pi t}{2}\right]$

3. $f(t) = -2\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin(nt)$

4. $f(t) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)t]}{(2n-1)^2}$

5.5 COMPLEX FOURIER SERIES

So far in our discussion, we expressed Fourier series in terms of sines and cosines. We are now ready to re-express a Fourier series as a series of complex exponentials. There are two reasons for this. First, in certain engineering and scientific applications of Fourier series, the expansion of a function in terms of complex exponentials results in coefficients of considerable simplicity and clarity. Second, these complex Fourier series point the way to the development of the Fourier transform in the next chapter.

We begin by introducing the variable $\omega_n = n\pi/L$, where $n = 0, \pm1, \pm2, \ldots$ Using Euler's formula we can replace the sine and cosine in the Fourier series by exponentials and find that

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\frac{a_n}{2}\left(e^{i\omega_n t} + e^{-i\omega_n t}\right) + \frac{b_n}{2i}\left(e^{i\omega_n t} - e^{-i\omega_n t}\right) \qquad \textbf{(5.5.1)}$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(\frac{a_n}{2} - \frac{b_n i}{2}\right)e^{i\omega_n t} + \left(\frac{a_n}{2} + \frac{b_n i}{2}\right)e^{-i\omega_n t}. \qquad \textbf{(5.5.2)}$$

If we define $c_n = \frac{1}{2}(a_n - ib_n)$, then

$$c_n = \tfrac{1}{2}(a_n - ib_n) = \frac{1}{2L}\int_\tau^{\tau+2L} f(t)[\cos(\omega_n t) - i\sin(\omega_n t)]\,dt = \frac{1}{2L}\int_\tau^{\tau+2L} f(t)e^{-i\omega_n t}dt. \tag{5.5.3}$$

Similarly, the complex conjugate of c_n, c_n^*, equals

$$c_n^* = \tfrac{1}{2}(a_n + ib_n) = \frac{1}{2L}\int_\tau^{\tau+2L} f(t)e^{i\omega_n t}dt. \tag{5.5.4}$$

To simplify Equation 5.5.2 we note that

$$\omega_{-n} = \frac{(-n)\pi}{L} = -\frac{n\pi}{L} = -\omega_n, \tag{5.5.5}$$

which yields the result that

$$c_{-n} = \frac{1}{2L}\int_\tau^{\tau+2L} f(t)e^{-i\omega_{-n} t}dt = \frac{1}{2L}\int_\tau^{\tau+2L} f(t)e^{i\omega_n t}dt = c_n^* \tag{5.5.6}$$

so that we can write Equation 5.5.2 as

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} c_n e^{i\omega_n t} + c_n^* e^{-i\omega_n t} = \frac{a_0}{2} + \sum_{n=1}^{\infty} c_n e^{i\omega_n t} + c_{-n} e^{-i\omega_n t}. \tag{5.5.7}$$

Letting $n = -m$ in the second summation on the right side of Equation 5.5.7,

$$\sum_{n=1}^{\infty} c_{-n} e^{-i\omega_n t} = \sum_{m=-1}^{-\infty} c_m e^{-i\omega_{-m} t} = \sum_{m=-\infty}^{-1} c_m e^{i\omega_m t} = \sum_{n=-\infty}^{-1} c_n e^{i\omega_n t}, \tag{5.5.8}$$

where we introduced $m = n$ into the last summation in Equation 5.5.8. Therefore,

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} c_n e^{i\omega_n t} + \sum_{n=-\infty}^{-1} c_n e^{i\omega_n t}. \tag{5.5.9}$$

On the other hand,

$$\frac{a_0}{2} = \frac{1}{2L}\int_\tau^{\tau+2L} f(t)\,dt = c_0 = c_0 e^{i\omega_0 t}, \tag{5.5.10}$$

because $\omega_0 = 0\pi/L = 0$. Thus, our final result is

$$\boxed{f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\omega_n t},} \tag{5.5.11}$$

where

$$\boxed{c_n = \frac{1}{2L}\int_\tau^{\tau+2L} f(t)e^{-i\omega_n t}\,dt} \tag{5.5.12}$$

and $n = 0, \pm 1, \pm 2, \ldots$. Note that even though c_n is generally complex, the summation Equation 5.5.11 always gives a *real-valued* function $f(t)$.

Just as we can represent the function $f(t)$ graphically by a plot of t against $f(t)$, we can plot c_n as a function of n, commonly called the frequency *spectrum*. Because c_n is generally complex, it is necessary to make two plots. Typically the plotted quantities are the amplitude spectra $|c_n|$ and the phase spectra φ_n, where φ_n is the phase of c_n. However, we could just as well plot the real and imaginary parts of c_n. Because n is an integer, these plots consist merely of a series of vertical lines representing the ordinates of the quantity $|c_n|$ or φ_n for each n. For this reason we refer to these plots as the *line spectra*.

Because $2c_n = a_n - ib_n$, the coefficients c_n for an even function will be purely real; the coefficients c_n for an odd function are purely imaginary. It is important to note that we lose the advantage of even and odd functions in the sense that we cannot just integrate over the interval 0 to L and then double the result. In the present case we have a line integral of a complex function along the real axis.

• Example 5.5.1

Let us find the complex Fourier series for

$$f(t) = \begin{cases} 1, & 0 < t < \pi, \\ -1, & -\pi < t < 0, \end{cases} \tag{5.5.13}$$

which has the periodicity $f(t + 2\pi) = f(t)$.

With $L = \pi$ and $\tau = -\pi$, $\omega_n = n\pi/L = n$. Therefore,

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{0} (-1) e^{-int}\, dt + \frac{1}{2\pi} \int_{0}^{\pi} (1) e^{-int}\, dt \tag{5.5.14}$$

$$= \frac{1}{2n\pi i} e^{-int} \bigg|_{-\pi}^{0} - \frac{1}{2n\pi i} e^{-int} \bigg|_{0}^{\pi} \tag{5.5.15}$$

$$= -\frac{i}{2n\pi} \left(1 - e^{n\pi i}\right) + \frac{i}{2n\pi} \left(e^{-n\pi i} - 1\right), \tag{5.5.16}$$

if $n \neq 0$. Because $e^{n\pi i} = \cos(n\pi) + i\sin(n\pi) = (-1)^n$ and $e^{-n\pi i} = \cos(-n\pi) + i\sin(-n\pi) = (-1)^n$, then

$$c_n = -\frac{i}{n\pi}[1 - (-1)^n] = \begin{cases} 0, & n \text{ even}, \\ -\frac{2i}{n\pi}, & n \text{ odd}, \end{cases} \tag{5.5.17}$$

with

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}. \tag{5.5.18}$$

In this particular problem we must treat the case $n = 0$ specially because Equation 5.5.15 is undefined for $n = 0$. In that case,

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{0} (-1)\, dt + \frac{1}{2\pi} \int_{0}^{\pi} (1)\, dt = \frac{1}{2\pi} (-t)\big|_{-\pi}^{0} + \frac{1}{2\pi} (t)\big|_{0}^{\pi} = 0. \tag{5.5.19}$$

Because $c_0 = 0$, we can write the expansion:

$$f(t) = -\frac{2i}{\pi} \sum_{m=-\infty}^{\infty} \frac{e^{(2m-1)it}}{2m - 1}, \tag{5.5.20}$$

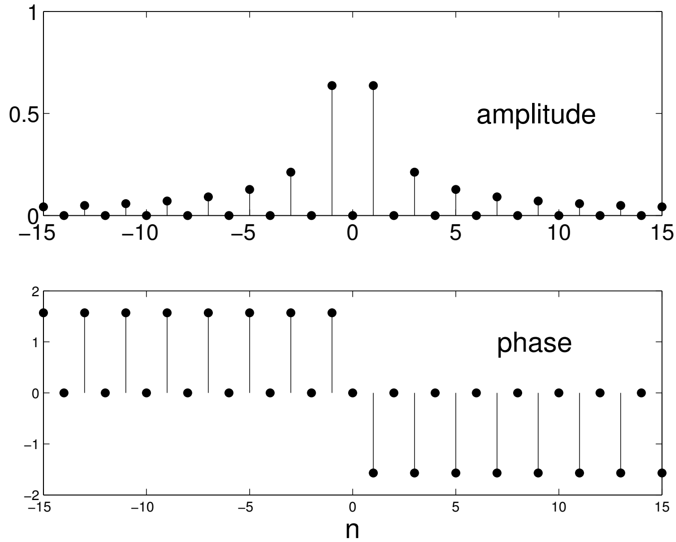


Figure 5.5.1: Amplitude and phase spectra for the function, Equation 5.5.13.

since we can write all odd integers as $2m-1$, where $m = 0, \pm1, \pm2, \pm3, \ldots$. Using the MATLAB script

```
max = 31; % total number of harmonics
mid = (max+1)/2; % in the array, location of c_0
for m = 1:max;
  n = m - mid; % compute value of harmonic
% compute complex Fourier coefficient c_n = (cnr,cni)
  if mod(n,2) == 0; cnr(m) = 0; cni(m) = 0; else;
  cnr(m) = 0; cni(m) = - 2/(pi*n); end;
end
nn=(1-mid):(max-mid); % create indices for x-axis
fzero=zeros(size(nn)); % create the zero line
clf % clear any figures
amplitude = sqrt(cnr.*cnr+cni.*cni);
phase = atan2(cni,cnr);
% plot amplitude of c_n
subplot(2,1,1), stem(nn,amplitude,'filled')
% label amplitude plot
text(6,0.75,'amplitude','FontSize',20)
subplot(2,1,2), stem(nn,phase,'filled') % plot phases of c_n
text(7,1,'phase','FontSize',20) % label phase plot
xlabel('n','Fontsize',20) % label x-axis,
```

we plot the amplitude and phase spectra for the function, Equation 5.5.13, as a function of n in Figure 5.5.1. □

• Example 5.5.2

The concept of Fourier series can be generalized to multivariable functions. Consider the function $f(x, y)$ defined over $0 < x < L$ and $0 < y < H$. Taking y constant, we have that

$$c_n(y) = \frac{1}{L} \int_0^L f(x,y) e^{-i\xi_n x} \, dx, \qquad \xi_n = \frac{2\pi n}{L}. \tag{5.5.21}$$

Similarly, holding ξ_n constant,

$$c_{nm} = \frac{1}{H} \int_0^H c_n(y) e^{-i\eta_m y} \, dy, \qquad \eta_m = \frac{2\pi m}{H}. \tag{5.5.22}$$

Therefore, the (complex) Fourier coefficient for the two-dimensional function $f(x,y)$ is

$$c_{nm} = \frac{1}{LH} \int_0^L \int_0^H f(x,y) e^{-i(\xi_n x + \eta_m y)} \, dx \, dy, \tag{5.5.23}$$

assuming that the integral exists.

To recover $f(x,y)$ given c_{nm}, we reverse the process of deriving c_{nm}. Starting with

$$c_n(y) = \sum_{m=-\infty}^{\infty} c_{nm} e^{i\eta_m y}, \tag{5.5.24}$$

we find that

$$f(x,y) = \sum_{n=-\infty}^{\infty} c_n(y) e^{i\xi_n x}. \tag{5.5.25}$$

Therefore,

$$f(x,y) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} c_{nm} e^{i(\xi_n x + \eta_m y)}. \tag{5.5.26}$$

Problems

Find the complex Fourier series for the following functions. Then use MATLAB to plot the corresponding spectra.

1. $f(t) = |t|, \quad -\pi \le t \le \pi$
2. $f(t) = e^t, \quad 0 < t < 2$
3. $f(t) = t, \quad 0 < t < 2$
4. $f(t) = t^2, \quad -\pi \le t \le \pi$
5. $f(t) = \begin{cases} 0, & -\pi/2 < t < 0 \\ 1, & 0 < t < \pi/2 \end{cases}$
6. $f(t) = t, \quad -1 < t < 1$

5.6 THE USE OF FOURIER SERIES IN THE SOLUTION OF ORDINARY DIFFERENTIAL EQUATIONS

An important application of Fourier series is the solution of ordinary differential equations. Structural engineers especially use this technique because the occupants of buildings and bridges often subject these structures to forcings that are periodic in nature.[17]

[17] Timoshenko, S. P., 1943: Theory of suspension bridges. Part II. *J. Franklin Inst.*, **235**, 327–349; Inglis, C. E., 1934: *A Mathematical Treatise on Vibrations in Railway Bridges*. Cambridge University Press, 203 pp.

• **Example 5.6.1**

Let us find the general solution to the ordinary differential equation

$$y'' + 9y = f(t), \tag{5.6.1}$$

where the forcing is

$$f(t) = |t|, \qquad -\pi \le t \le \pi, \qquad f(t + 2\pi) = f(t). \tag{5.6.2}$$

This equation represents an oscillator forced by a driver whose displacement is the saw-tooth function.

We begin by replacing the function $f(t)$ by its Fourier series representation because the forcing function is periodic. The advantage of expressing $f(t)$ as a Fourier series is its validity for any time t. The alternative would be to construct a solution over each interval $n\pi < t < (n+1)\pi$ and then piece together the final solution assuming that the solution and its first derivative are continuous at each junction $t = n\pi$. Because the function is an even function, all of the sine terms vanish and the Fourier series is

$$|t| = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)t]}{(2n-1)^2}. \tag{5.6.3}$$

Next, we note that the general solution consists of the complementary solution, which equals

$$y_H(t) = A\cos(3t) + B\sin(3t), \tag{5.6.4}$$

and the particular solution $y_p(t)$, which satisfies the differential equation

$$y_p'' + 9y_p = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)t]}{(2n-1)^2}. \tag{5.6.5}$$

To determine this particular solution, we write Equation 5.6.5 as

$$y_p'' + 9y_p = \frac{\pi}{2} - \frac{4}{\pi}\cos(t) - \frac{4}{9\pi}\cos(3t) - \frac{4}{25\pi}\cos(5t) - \cdots. \tag{5.6.6}$$

By the method of undetermined coefficients, we guess the particular solution:

$$y_p(t) = \frac{a_0}{2} + a_1\cos(t) + b_1\sin(t) + a_2\cos(3t) + b_2\sin(3t) + \cdots \tag{5.6.7}$$

or

$$y_p(t) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n\cos[(2n-1)t] + b_n\sin[(2n-1)t]. \tag{5.6.8}$$

Because

$$y_p''(t) = \sum_{n=1}^{\infty} -(2n-1)^2\{a_n\cos[(2n-1)t] + b_n\sin[(2n-1)t]\}, \tag{5.6.9}$$

$$\sum_{n=1}^{\infty} -(2n-1)^2\{a_n\cos[(2n-1)t] + b_n\sin[(2n-1)t]\} \tag{5.6.10}$$

$$+ \tfrac{9}{2}a_0 + 9\sum_{n=1}^{\infty} a_n\cos[(2n-1)t] + b_n\sin[(2n-1)t] = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)t]}{(2n-1)^2},$$

or

$$\frac{9a_0}{2} - \frac{\pi}{2} + \sum_{n=1}^{\infty} \left\{ [9 - (2n-1)^2] a_n + \frac{4}{\pi(2n-1)^2} \right\} \cos[(2n-1)t]$$
$$+ \sum_{n=1}^{\infty} [9 - (2n-1)^2] b_n \sin[(2n-1)t] = 0. \tag{5.6.11}$$

Because Equation 5.6.11 must hold true for any time, each harmonic must vanish separately and

$$a_0 = \frac{\pi}{9}, \qquad a_n = -\frac{4}{\pi(2n-1)^2[9-(2n-1)^2]} \tag{5.6.12}$$

and $b_n = 0$. All of the coefficients a_n are finite except for $n = 2$, where a_2 becomes undefined. This coefficient is undefined because the harmonic $\cos(3t)$ in the forcing function resonates with the natural mode of the system.

Let us review our analysis to date. We found that each harmonic in the forcing function yields a corresponding harmonic in the particular solution, Equation 5.6.8. The only difficulty arises with the harmonic $n = 2$. Although our particular solution is not correct because it contains $\cos(3t)$, we suspect that if we remove that term then the remaining harmonic solutions are correct. The problem is linear, and difficulties with one harmonic term should not affect other harmonics. But how shall we deal with the $\cos(3t)$ term in the forcing function? Let us denote that particular solution by $Y(t)$ and modify our particular solution as follows:

$$y_p(t) = \tfrac{1}{2}a_0 + a_1\cos(t) + Y(t) + a_3\cos(5t) + \cdots. \tag{5.6.13}$$

Substituting this solution into the differential equation and simplifying, everything cancels except

$$Y'' + 9Y = -\frac{4}{9\pi}\cos(3t). \tag{5.6.14}$$

The solution of this equation by the method of undetermined coefficients is

$$Y(t) = -\frac{2}{27\pi} t \sin(3t). \tag{5.6.15}$$

This term, called a *secular term*, is the most important one in the solution. While the other terms merely represent simple oscillatory motion, the term $t\sin(3t)$ grows linearly with time and eventually becomes the dominant term in the series. Consequently, the general solution equals the complementary plus the particular solution, or

$$y(t) = A\cos(3t) + B\sin(3t) + \frac{\pi}{18} - \frac{2}{27\pi}t\sin(3t) - \frac{4}{\pi}\sum_{\substack{n=1\\ n\neq 2}}^{\infty} \frac{\cos[(2n-1)t]}{(2n-1)^2[9-(2n-1)^2]}. \tag{5.6.16}$$

□

• Example 5.6.2

Let us redo the previous problem only using complex Fourier series. That is, let us find the general solution to the ordinary differential equation

$$y'' + 9y = \frac{\pi}{2} - \frac{2}{\pi}\sum_{n=-\infty}^{\infty} \frac{e^{i(2n-1)t}}{(2n-1)^2}. \tag{5.6.17}$$

From the method of undetermined coefficients we guess the particular solution for Equation 5.6.17 to be

$$y_p(t) = c_0 + \sum_{n=-\infty}^{\infty} c_n e^{i(2n-1)t}. \tag{5.6.18}$$

Then

$$y_p''(t) = \sum_{n=-\infty}^{\infty} -(2n-1)^2 c_n e^{i(2n-1)t}. \tag{5.6.19}$$

Substituting Equation 5.6.18 and Equation 5.6.19 into Equation 5.6.17,

$$9c_0 + \sum_{n=-\infty}^{\infty} [9-(2n-1)^2] c_n e^{i(2n-1)t} = \frac{\pi}{2} - \frac{2}{\pi} \sum_{n=-\infty}^{\infty} \frac{e^{i(2n-1)t}}{(2n-1)^2}. \tag{5.6.20}$$

Because Equation 5.6.20 must be true for any t,

$$c_0 = \frac{\pi}{18}, \quad \text{and} \quad c_n = \frac{2}{\pi(2n-1)^2[(2n-1)^2-9]}. \tag{5.6.21}$$

Therefore,

$$y_p(t) = \frac{\pi}{18} + \frac{2}{\pi} \sum_{n=-\infty}^{\infty} \frac{e^{i(2n-1)t}}{(2n-1)^2[(2n-1)^2-9]} e^{i(2n-1)t}. \tag{5.6.22}$$

However, there is a problem when $n=-1$ and $n=2$. Therefore, we modify Equation 5.6.22 to read

$$y_p(t) = \frac{\pi}{18} + c_2 t e^{3it} + c_{-1} t e^{-3it} + \frac{2}{\pi} \sum_{\substack{n=-\infty \\ n\neq -1,2}}^{\infty} \frac{e^{i(2n-1)t}}{(2n-1)^2[(2n-1)^2-9]} e^{i(2n-1)t}. \tag{5.6.23}$$

Introducing Equation 5.6.23 into Equation 5.6.17 and simplifying,

$$c_2 = -\frac{1}{27\pi i}, \qquad \text{and} \qquad c_{-1} = -\frac{1}{27\pi i}. \tag{5.6.24}$$

The general solution is then

$$y(t) = Ae^{3it} + Be^{-3it} + \frac{\pi}{18} - \frac{te^{3it}}{27\pi i} + \frac{te^{-3it}}{27\pi i} + \frac{2}{\pi} \sum_{\substack{n=-\infty \\ n\neq -1,2}}^{\infty} \frac{e^{i(2n-1)t}}{(2n-1)^2[(2n-1)^2-9]}. \tag{5.6.25}$$

The first two terms on the right side of Equation 5.6.25 represent the complementary solution. Although this expansion is equivalent to Equation 5.6.16, we have all of the advantages of dealing with exponentials rather than sines and cosines. These advantages include ease of differentiation and integration, and writing the series in terms of amplitude and phase. □

• Example 5.6.3: Temperature within a spinning satellite

In the design of artificial satellites, it is important to determine the temperature distribution on the spacecraft's surface. An interesting special case is the temperature fluctuation in the skin due to the spinning of the vehicle. If the craft is thin-walled so that there is no radial dependence, Hrycak[18] showed that he could approximate the nondimensional temperature field at the equator of the rotating satellite by

$$\frac{d^2T}{d\eta^2} + b\frac{dT}{d\eta} - c\left(T - \frac{3}{4}\right) = -\frac{\pi c}{4}\frac{F(\eta) + \beta/4}{1 + \pi\beta/4}, \tag{5.6.26}$$

where

$$b = 4\pi^2 r^2 f/a, \quad c = \frac{16\pi S}{\gamma T_\infty}\left(1 + \frac{\pi\beta}{4}\right), \quad T_\infty = \left(\frac{S}{\pi\sigma\epsilon}\right)^{1/4}\left(\frac{1 + \pi\beta/4}{1 + \beta}\right)^{1/4}, \tag{5.6.27}$$

$$F(\eta) = \begin{cases} \cos(2\pi\eta), & 0 \le \eta \le \frac{1}{4}, \\ 0, & \frac{1}{4} \le \eta \le \frac{3}{4}, \\ \cos(2\pi\eta), & \frac{3}{4} \le \eta \le 1, \end{cases} \tag{5.6.28}$$

a is the thermal diffusivity of the shell, f is the rate of spin, r is the radius of the spacecraft, S is the net direct solar heating, β is the ratio of the emissivity of the interior shell to the emissivity of the exterior surface, ϵ is the overall emissivity of the exterior surface, γ is the satellite's skin conductance, and σ is the Stefan-Boltzmann constant. The independent variable η is the longitude along the equator with the effect of rotation subtracted out ($2\pi\eta = \varphi - 2\pi ft$). The reference temperature T_∞ equals the temperature that the spacecraft would have if it spun with infinite angular speed so that the solar heating would be uniform around the craft. We nondimensionalized the temperature with respect to T_∞.

We begin by introducing the new variables

$$y = T - \frac{3}{4} - \frac{\pi\beta}{16 + 4\pi\beta}, \quad \nu_0 = \frac{2\pi^2 r^2 f}{a\rho_0}, \quad A_0 = -\frac{\pi\rho^2}{4 + \pi\beta} \tag{5.6.29}$$

and $\rho_0^2 = c$ so that Equation 5.6.26 becomes

$$\frac{d^2y}{d\eta^2} + 2\rho_0\nu_0\frac{dy}{d\eta} - \rho_0^2 y = A_0 F(\eta). \tag{5.6.30}$$

Next, we expand $F(\eta)$ as a Fourier series because it is a periodic function of period 1. Because it is an even function,

$$f(\eta) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(2n\pi\eta), \tag{5.6.31}$$

where

$$a_0 = \frac{1}{1/2}\int_0^{1/4} \cos(2\pi x)\,dx + \frac{1}{1/2}\int_{3/4}^{1} \cos(2\pi x)\,dx = \frac{2}{\pi}, \tag{5.6.32}$$

$$a_1 = \frac{1}{1/2}\int_0^{1/4} \cos^2(2\pi x)\,dx + \frac{1}{1/2}\int_{3/4}^{1} \cos^2(2\pi x)\,dx = \frac{1}{2} \tag{5.6.33}$$

---

[18] From Hrycak, P., 1963: Temperature distribution in a spinning spherical space vehicle. *AIAA J.*, **1**, 96–99. Reprinted with permission of the American Institute of Aeronautics and Astronautics.

and

$$a_n = \frac{1}{1/2}\int_0^{1/4} \cos(2\pi x)\cos(2n\pi x)\,dx + \frac{1}{1/2}\int_{3/4}^1 \cos(2\pi x)\cos(2n\pi x)\,dx \tag{5.6.34}$$

$$= -\frac{2(-1)^n}{\pi(n^2-1)}\cos\left(\frac{n\pi}{2}\right), \tag{5.6.35}$$

if $n \geq 2$. Therefore,

$$f(\eta) = \frac{1}{\pi} + \frac{1}{2}\cos(2\pi\eta) - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{4n^2-1}\cos(4n\pi\eta). \tag{5.6.36}$$

From the method of undetermined coefficients, the particular solution is

$$y_p(\eta) = \tfrac{1}{2}a_0 + a_1\cos(2\pi\eta) + b_1\sin(2\pi\eta) + \sum_{n=1}^{\infty} a_{2n}\cos(4n\pi\eta) + b_{2n}\sin(4n\pi\eta), \tag{5.6.37}$$

which yields

$$y_p'(\eta) = -2\pi a_1\sin(2\pi\eta) + 2\pi b_1\cos(2\pi\eta) + \sum_{n=1}^{\infty}[-4n\pi a_{2n}\sin(4n\pi\eta) + 4n\pi b_{2n}\cos(4n\pi\eta)], \tag{5.6.38}$$

and

$$y_p''(\eta) = -4\pi^2[a_1\cos(2\pi\eta) + b_1\sin(2\pi\eta)] - \sum_{n=1}^{\infty} 16n^2\pi^2[a_{2n}\cos(4n\pi\eta) + b_{2n}\sin(4n\pi\eta)]. \tag{5.6.39}$$

Substituting into Equation 5.6.30,

$$\begin{aligned} -\frac{1}{2}\rho_0^2 a_0 - \frac{A_0}{\pi} &+ \left(-4\pi^2 a_1 + 4\pi\rho_0\nu_0 b_1 - \rho_0^2 a_1 - \frac{A_0}{2}\right)\cos(2\pi\eta) \\ &+ \left(-4\pi^2 b_1 - 4\pi\rho_0\nu_0 a_1 - \rho_0^2 b_1\right)\sin(2\pi\eta) \\ &+ \sum_{n=1}^{\infty}\left[-16n^2\pi^2 a_{2n} + 8n\pi\rho_0\nu_0 b_{2n} - \rho_0^2 a_{2n} + \frac{2A_0(-1)^n}{\pi(4n^2-1)}\right]\cos(4n\pi\eta) \\ &+ \sum_{n=1}^{\infty}\left(-16n^2\pi^2 b_{2n} - 8n\pi\rho_0\nu_0 a_{2n} - \rho_0^2 b_{2n}\right)\sin(4n\pi\eta) = 0. \end{aligned} \tag{5.6.40}$$

To satisfy Equation 5.6.40 for any η, we set

$$a_0 = -\frac{2A_0}{\pi\rho_0^2}, \tag{5.6.41}$$

$$-(4\pi^2 + \rho_0^2)a_1 + 4\pi\rho_0\nu_0 b_1 = \frac{A_0}{2}, \tag{5.6.42}$$

$$4\pi\rho_0\nu_0 a_1 + (4\pi^2 + \rho_0^2)b_1 = 0, \tag{5.6.43}$$

$$(16n^2\pi^2 + \rho_0^2)a_{2n} - 8n\pi\rho_0\nu_0 b_{2n} = \frac{2A_0(-1)^n}{\pi(4n^2-1)}, \tag{5.6.44}$$

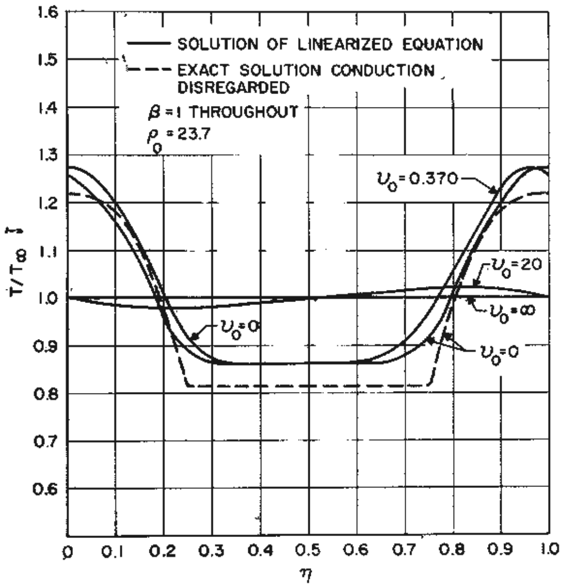


Figure 5.6.1: Temperature distribution along the equator of a spinning spherical satellite. (From Hrycak, P., 1963: Temperature distribution in a spinning spherical space vehicle. *AIAA J.*, **1**, 97. ©1963 AIAA, reprinted with permission.)

and

$$8n\pi\rho_0\nu_0 a_{2n} + (16n^2\pi^2 + \rho_0^2)b_{2n} = 0, \tag{5.6.45}$$

or

$$[16\pi^2\rho_0^2\nu_0^2 + (4\pi^2 + \rho_0^2)^2]a_1 = -\frac{(4\pi^2 + \rho_0^2)A_0}{2}, \tag{5.6.46}$$

$$[16\pi^2\rho_0^2\nu_0^2 + (4\pi^2 + \rho_0^2)^2]b_1 = 2\pi\rho_0\nu_0 A_0, \tag{5.6.47}$$

$$[64n^2\pi^2\rho_0^2\nu_0^2 + (16n^2\pi^2 + \rho_0^2)^2]a_{2n} = \frac{2A_0(-1)^n(16n^2\pi^2 + \rho_0^2)}{\pi(4n^2 - 1)}, \tag{5.6.48}$$

and

$$[64n^2\pi^2\rho_0^2\nu_0^2 + (16n^2\pi^2 + \rho_0^2)^2]b_{2n} = -\frac{16(-1)^n\rho_0\nu_0 n A_0}{4n^2 - 1}. \tag{5.6.49}$$

Substituting for a_0, a_1, b_1, a_{2n}, and b_{2n}, the particular solution is

$$\begin{aligned} y_p(\eta) = &-\frac{A_0}{\pi\rho_0^2} - \frac{(4\pi^2 + \rho_0^2)A_0\cos(2\pi\eta)}{2[(4\pi^2 + \rho_0^2)^2 + 16\pi^2\rho_0^2\nu_0^2]} + \frac{2\pi\rho_0\nu_0 A_0\sin(2\pi\eta)}{(4\pi^2 + \rho_0^2)^2 + 16\pi^2\rho_0^2\nu_0^2} \\ &+ \frac{2A_0}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n(16n^2\pi^2 + \rho_0^2)\cos(2n\pi\eta)}{(4n^2 - 1)[64n^2\pi^2\rho_0^2\nu_0^2 + (16n^2\pi^2 + \rho_0^2)^2]} \\ &- 16\rho_0\nu_0 A_0\sum_{n=1}^{\infty}\frac{(-1)^n n\sin(2n\pi\eta)}{(4n^2 - 1)[64n^2\pi^2\rho_0^2\nu_0^2 + (16n^2\pi^2 + \rho_0^2)^2]}. \end{aligned} \tag{5.6.50}$$

Figure 5.6.1 is from Hrycak's paper and shows the variation of the nondimensional temperature as a function of η for the spinning rate ν_0. The other parameters are typical of a satellite with aluminum skin and fully covered with glass-protected solar cells. As a check on the solution, we show the temperature field (the dashed line) of a nonrotating satellite where we neglect the effects of conduction and only radiation occurs. The difference between the $\nu_0 = 0$ solid and dashed lines arises primarily due to the *linearization* of the nonlinear radiation boundary condition during the derivation of the governing equations.

Problems

Solve the following ordinary differential equations by Fourier series if the forcing is given by the periodic function

$$f(t) = \begin{cases} 1, & 0 < t < \pi, \\ 0, & \pi < t < 2\pi, \end{cases}$$

and $f(t) = f(t + 2\pi)$:

1. $y'' - y = f(t)$,
2. $y'' + y = f(t)$,
3. $y'' - 3y' + 2y = f(t)$.

Solve the following ordinary differential equations by *complex* Fourier series if the forcing is given by the periodic function

$$f(t) = |t|, \qquad -\pi \le t \le \pi,$$

and $f(t) = f(t + 2\pi)$:

4. $y'' - y = f(t)$,
5. $y'' + 4y = f(t)$.

6. An object radiating into its nocturnal surrounding has a temperature $y(t)$ governed by the equation[19]

$$\frac{dy}{dt} + ay = A_0 + \sum_{n=1}^{\infty} A_n \cos(n\omega t) + B_n \sin(n\omega t),$$

where the constant a is the heat loss coefficient and the Fourier series describes the temporal variation of the atmospheric air temperature and the effective sky temperature. If $y(0) = T_0$, find $y(t)$.

7. The equation that governs the charge q on the capacitor of an LRC electrical circuit is

$$q'' + 2\alpha q' + \omega^2 q = \omega^2 E,$$

where $\alpha = R/(2L)$, $\omega^2 = 1/(LC)$, R denotes resistance, C denotes capacitance, L denotes the inductance, and E is the electromotive force driving the circuit. If E is given by

$$E = \sum_{n=-\infty}^{\infty} \varphi_n e^{in\omega_0 t},$$

find $q(t)$.

8. Use Fourier series to find the *particular* solution[20] of the ordinary differential equation

$$y''(x) + k^2 y(x) = -k^2 V_L \begin{cases} 1, & 0 < x < \lambda/4, \\ -1, & \lambda/4 < x < \lambda/2, \end{cases}$$

---

[19] See Sodha, M. S., 1982: Transient radiative cooling. *Solar Energy*, **28**, 541.

[20] See Chabert, P., J. L. Raimbault, J. M. Rax, and M. A. Lieberman, 2004: Self-consistent nonlinear transmission line model of standing wave effects in a capacitive discharge. *Phys. Plasmas*, **11**, 1775–1785.

where k, V_L and λ are constants. Extend the forcing function as an even function into the interval $(-\lambda/2, 0)$.

5.7 FINITE FOURIER SERIES

In many applications we must construct a Fourier series from values given by data or a graph. Unlike the situation with analytic formulas where we have an infinite number of data points and, consequently, an infinite number of terms in the Fourier series, the Fourier series contains a finite number of sines and cosines where the number of coefficients equals the number of data points.

Assuming that these series are useful, the next question is how do we find the Fourier coefficients? We could compute them by numerically integrating Equation 5.1.6. However, the results would suffer from the truncation errors that afflict all numerical schemes. On the other hand, we can avoid this problem if we again employ the orthogonality properties of sines and cosines, now in their discrete form. Just as in the case of conventional Fourier series, we can use these properties to derive formulas for computing the Fourier coefficients. These results will be *exact* except for roundoff errors.

We start by deriving some preliminary results. Let us define $x_m = mP/(2N)$. Then, if k is an integer,

$$\sum_{m=0}^{2N-1} \exp\left(\frac{2\pi i k x_m}{P}\right) = \sum_{m=0}^{2N-1} \exp\left(\frac{km\pi i}{N}\right) = \sum_{m=0}^{2N-1} r^m = \begin{cases} \frac{1-r^{2N}}{1-r} = 0, & r \neq 1, \\ 2N, & r = 1, \end{cases} \tag{5.7.1}$$

because $r^{2N} = \exp(2\pi k i) = 1$ if $r \neq 1$. If $r = 1$, then the sum consists of $2N$ terms, each of which equals one. The condition $r = 1$ corresponds to $k = 0, \pm 2N, \pm 4N, \ldots$. Taking the real and imaginary part of Equation 5.7.1,

$$\sum_{m=0}^{2N-1} \cos\left(\frac{2\pi k x_m}{P}\right) = \begin{cases} 0, & k \neq 0, \pm 2N, \pm 4N, \ldots, \\ 2N, & k = 0, \pm 2N, \pm 4N, \ldots, \end{cases} \tag{5.7.2}$$

and

$$\sum_{m=0}^{2N-1} \sin\left(\frac{2\pi k x_m}{P}\right) = 0 \tag{5.7.3}$$

for all k.

Consider now the following sum:

$$\sum_{m=0}^{2N-1} \cos\left(\frac{2\pi k x_m}{P}\right) \cos\left(\frac{2\pi j x_m}{P}\right) = \frac{1}{2} \sum_{m=0}^{2N-1} \left\{ \cos\left[\frac{2\pi(k+j)x_m}{P}\right] + \cos\left[\frac{2\pi(k-j)x_m}{P}\right] \right\} \tag{5.7.4}$$

$$= \begin{cases} 0, & |k-j| \text{ and } |k+m| \neq 0, 2N, 4N, \ldots, \\ N, & |k-j| \text{ or } |k+m| \neq 0, 2N, 4N, \ldots, \\ 2N, & |k-j| \text{ and } |k+m| = 0, 2N, 4N, \ldots. \end{cases} \tag{5.7.5}$$

Let us simplify the right side of Equation 5.7.5 by restricting ourselves to $k+j$ lying between 0 to $2N$. This is permissible because of the periodic nature of this equation. If $k+j = 0$, $k = j = 0$; if $k+j = 2N$, $k = j = N$. In either case, $k-j = 0$ and the right side of Equation 5.7.5 equals $2N$. Consider now the case $k \neq j$. Then $k+j \neq 0$ or $2N$ and $k-j \neq 0$ or $2N$.

The right side of this equaton must equal 0. Finally, if $k = j \neq 0$ or N, then $k + j \neq 0$ or $2N$ but $k - j = 0$ and the right side of this equation equals N. In summary,

$$\sum_{m=0}^{2N-1} \cos\left(\frac{2\pi k x_m}{P}\right) \cos\left(\frac{2\pi j x_m}{P}\right) = \begin{cases} 0, & k \neq j \\ N, & k = j \neq 0, N \\ 2N, & k = j = 0, N. \end{cases} \tag{5.7.6}$$

In a similar manner,

$$\sum_{m=0}^{2N-1} \cos\left(\frac{2\pi k x_m}{P}\right) \sin\left(\frac{2\pi j x_m}{P}\right) = 0 \tag{5.7.7}$$

for all k and j and

$$\sum_{m=0}^{2N-1} \sin\left(\frac{2\pi k x_m}{P}\right) \sin\left(\frac{2\pi j x_m}{P}\right) = \begin{cases} 0, & k \neq j \\ N, & k = j \neq 0, N, \\ 0, & k = j = 0, N. \end{cases} \tag{5.7.8}$$

Armed with these equations we are ready to find the coefficients A_n and B_n of the finite Fourier series,

$$f(x) = \frac{A_0}{2} + \sum_{k=1}^{N-1} \left[A_k \cos\left(\frac{2\pi k x}{P}\right) + B_k \sin\left(\frac{2\pi k x}{P}\right) \right] + \frac{A_N}{2} \cos\left(\frac{2\pi N x}{P}\right), \tag{5.7.9}$$

where we have $2N$ data points and now define P as the period of the function.

To find A_k we proceed as before and multiply Equation 5.7.9 by $\cos(2\pi j x\ /P)$ (j may take on values from 0 to N) and sum from 0 to $2N - 1$. At the point $x = x_m$,

$$\begin{aligned} \sum_{m=0}^{2N-1} f(x_m) \cos\left(\frac{2\pi j}{P} x_m\right) &= \frac{A_0}{2} \sum_{m=0}^{2N-1} \cos\left(\frac{2\pi j}{P} x_m\right) \\ &+ \sum_{k=1}^{N-1} A_k \sum_{m=0}^{2N-1} \cos\left(\frac{2\pi k}{P} x_m\right) \cos\left(\frac{2\pi j}{P} x_m\right) \\ &+ \sum_{k=1}^{N-1} B_k \sum_{m=0}^{2N-1} \sin\left(\frac{2\pi k}{P} x_m\right) \cos\left(\frac{2\pi j}{P} x_m\right) \\ &+ \frac{A_N}{2} \sum_{m=0}^{2N-1} \cos\left(\frac{2\pi N}{P} x_m\right) \cos\left(\frac{2\pi j}{P} x_m\right). \end{aligned} \tag{5.7.10}$$

If $j \neq 0$ or N, then the first summation on the right side vanishes by Equation 5.7.2, the third by Equation 5.7.8, and the fourth by Equation 5.7.6. The second summation does *not* vanish if $k = j$ and equals N. Similar considerations lead to the formulas for the calculation of A_k and B_k:

$$A_k = \frac{1}{N} \sum_{m=0}^{2N-1} f(x_m) \cos\left(\frac{2\pi k}{P} x_m\right), \qquad k = 0, 1, 2, \ldots, N, \tag{5.7.11}$$

and

$$B_k = \frac{1}{N} \sum_{m=0}^{2N-1} f(x_m) \sin\left(\frac{2\pi k}{P} x_m\right), \qquad k = 1, 2, \ldots, N - 1. \tag{5.7.12}$$

Table 5.7.1: The Depth of Water in the Harbor at Buffalo, NY (Minus the Low-Water Datum of 568.8 ft) on the 15$^{\text{th}}$ Day of Each Month During 1977

mo	n	depth	mo	n	depth	mo	n	depth
Jan	1	1.61	May	5	3.16	Sep	9	2.42
Feb	2	1.57	Jun	6	2.95	Oct	10	2.95
Mar	3	2.01	Jul	7	3.10	Nov	11	2.74
Apr	4	2.68	Aug	8	2.90	Dec	12	2.63

If there are $2N+1$ data points and $f(x_0) = f(x_{2N})$, then Equation 5.7.11 and Equation 5.7.12 are still valid and we need only consider the first $2N$ points. If $f(x_0) \neq f(x_{2N})$, we can still use our formulas if we require that the endpoints have the value of $[f(x_0) + f(x_{2N})]/2$. In this case the formulas for the coefficients A_k and B_k are

$$A_k = \frac{1}{N}\left[\frac{f(x_0) + f(x_{2N})}{2} + \sum_{m=1}^{2N-1} f(x_m)\cos\left(\frac{2\pi k}{P}x_m\right)\right], \tag{5.7.13}$$

where $k = 0, 1, 2, \ldots, N$, and

$$B_k = \frac{1}{N}\sum_{m=1}^{2N-1} f(x_m)\sin\left(\frac{2\pi k}{P}x_m\right), \tag{5.7.14}$$

where $k = 1, 2, \ldots, N-1$.

It is important to note that $2N$ data points yield $2N$ Fourier coefficients A_k and B_k. Consequently our sampling frequency will always limit the amount of information, whether in the form of data points or Fourier coefficients. It might be argued that from the Fourier series representation of $f(t)$ we could find the value of $f(t)$ for any given t, which is more than we can do with the data alone. This is not true. Although we can calculate $f(t)$ at any t using the finite Fourier series, the values may or may not be correct since the constraint on the finite Fourier series is that the series must fit the data in a least-squared sense. Despite the limitations imposed by only having a finite number of Fourier coefficients, the Fourier analysis of finite data sets yields valuable physical insights into the processes governing many physical systems.

• Example 5.7.1: Water depth at Buffalo, NY

Each entry[21] in Table 5.7.1 gives the observed depth of water at Buffalo, NY (minus the low-water datum of 568.6 ft) on the 15$^{\text{th}}$ of the corresponding month during 1977. Assuming that the water level is a periodic function of 1 year, and that we took the observations at equal intervals, let us construct a finite Fourier series from these data. This corresponds to computing the Fourier coefficients $A_0, A_1, \ldots, A_6, B_1, \ldots, B_5$, which give the mean level and harmonic fluctuations of the depth of water, the harmonics having the periods 12 months, 6 months, 4 months, and so forth.

[21] National Ocean Survey, 1977: *Great Lakes Water Level, 1977, Daily and Monthly Average Water Surface Elevations*. National Oceanic and Atmospheric Administration.

In this problem, P equals 12 months, $N = P/2 = 6$ mo, and $x_m = mP/(2N) = m(12 \text{ mo})/12 \text{ mo} = m$. That is, there should be a data point for each month. From Equation 5.7.11 and Equation 5.7.12,

$$A_k = \frac{1}{6}\sum_{m=0}^{11} f(x_m)\cos\left(\frac{mk\pi}{6}\right), \quad k = 0, 1, 2, 3, 4, 5, 6, \tag{5.7.15}$$

and

$$B_k = \frac{1}{6}\sum_{m=0}^{11} f(x_m)\sin\left(\frac{mk\pi}{6}\right), \quad k = 1, 2, 3, 4, 5. \tag{5.7.16}$$

Substituting the data into these equations yields

A_0	= twice the mean level			= +5.120 ft
A_1	= harmonic component with a period of	12	mo	= −0.566 ft
B_1	= harmonic component with a period of	12	mo	= −0.128 ft
A_2	= harmonic component with a period of	6	mo	= −0.177 ft
B_2	= harmonic component with a period of	6	mo	= −0.372 ft
A_3	= harmonic component with a period of	4	mo	= −0.110 ft
B_3	= harmonic component with a period of	4	mo	= −0.123 ft
A_4	= harmonic component with a period of	3	mo	= +0.025 ft
B_4	= harmonic component with a period of	3	mo	= +0.052 ft
A_5	= harmonic component with a period of	2.4	mo	= −0.079 ft
B_5	= harmonic component with a period of	2.4	mo	= −0.131 ft
A_6	= harmonic component with a period of	2	mo	= −0.107 ft

Figure 5.7.1 is a plot of our results using Equation 5.7.9. Note that when we include all of the harmonic terms, the finite Fourier series fits the data points exactly. The values given by the series at points between the data points may be right or they may not. To illustrate this, we also plotted the values for the first of each month. Sometimes the values given by the Fourier series and these intermediate data points are quite different.

Let us now examine our results in terms of various physical processes. In the long run the depth of water in the harbor at Buffalo, NY depends upon the three-way balance between precipitation, evaporation, and inflow-outflow of any rivers. Because the inflow and outflow of the rivers depends strongly upon precipitation, and evaporation is of secondary importance, the water level should correlate with the precipitation rate. It is well known that more precipitation falls during the warmer months rather than the colder months. The large amplitude of the Fourier coefficient A_1 and B_1, corresponding to the annual cycle ($k = 1$), reflects this.

Another important term in the harmonic analysis corresponds to the semiannual cycle ($k = 2$). During the winter months around Lake Ontario, precipitation falls as snow. Therefore, the inflow from rivers is greatly reduced. When spring comes, the snow and ice melt and a jump in the water level occurs. Because the second harmonic gives periodic variations associated with seasonal variations, this harmonic is absolutely necessary if we want to get the correct answer while the higher harmonics do not represent any specific physical process. □

• **Example 5.7.2: Numerical computation of Fourier coefficients**

At the beginning of this chapter, we showed how you could compute the Fourier coefficients a_0, a_n, and b_n from Equation 5.1.6 given a function $f(t)$. All of this assumed that you

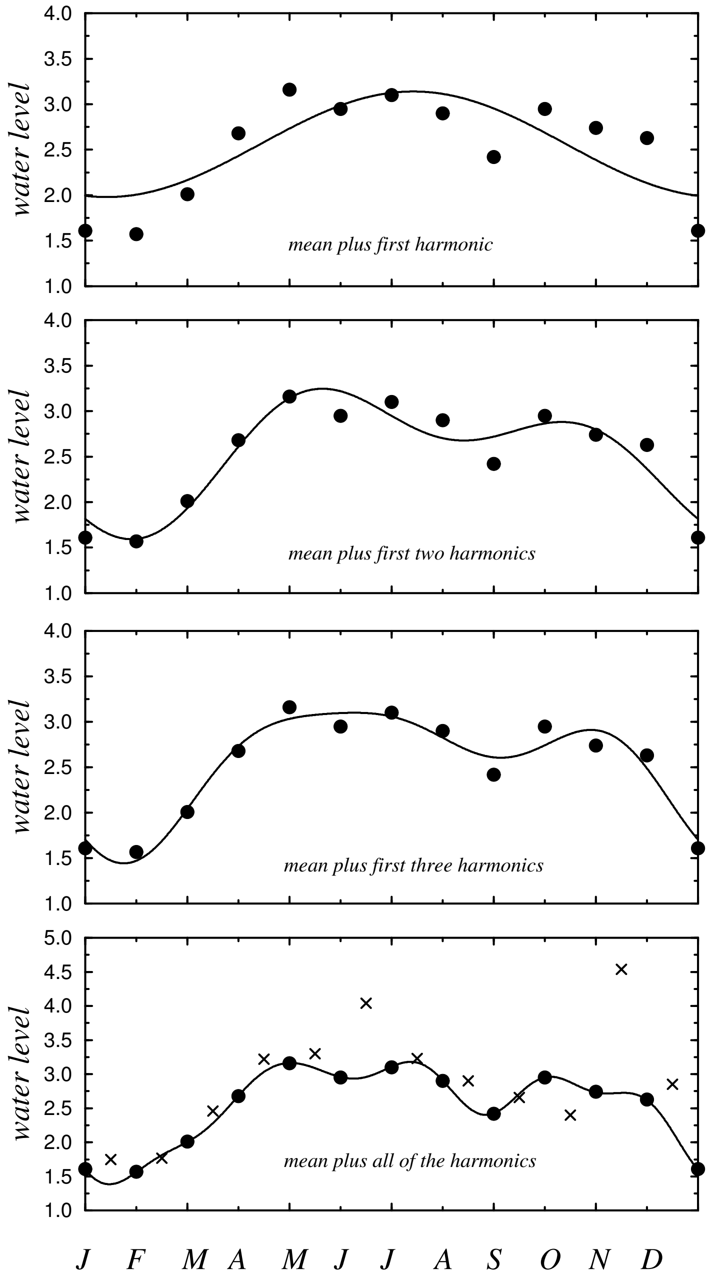


Figure 5.7.1: Partial sums of the finite Fourier series for the depth of water in the harbor of Buffalo, NY during 1977. Circles indicate observations on the 15^{th} of the month; crosses are observations on the first.

could carry out the integrations. What do you do if you cannot perform the integrations? The obvious solution is perform it numerically. In this section we showed that the best approximation to Equation 5.1.6 is given by Equation 5.7.11 and Equation 5.7.12. In the case when we have $f(t)$ this is still true but we may choose N as large as necessary to obtain the desired number of Fourier coefficients.

To illustrate this we have redone Example 5.1.1 and plotted the exact (analytic) and

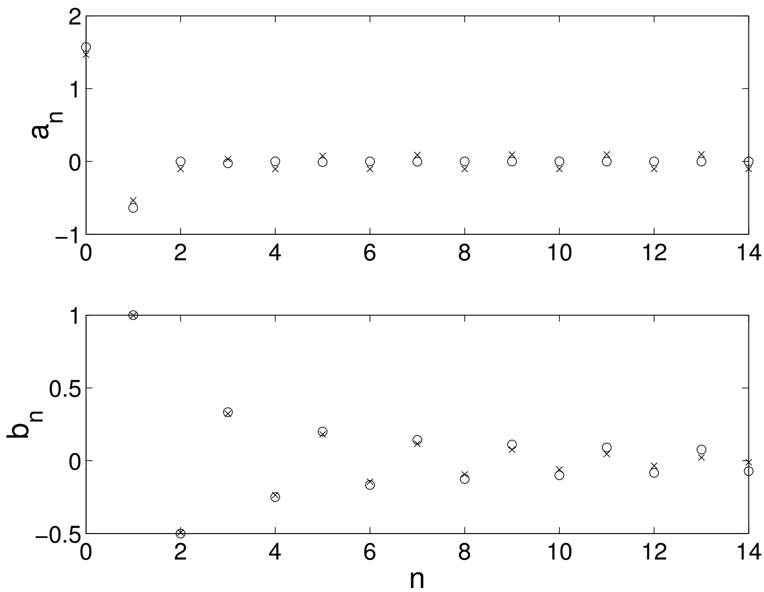


Figure 5.7.2: The computation of Fourier coefficients using a finite Fourier series when $f(t)$ is given by Equation 5.1.8. The circles give a_n and b_n as computed from Equation 5.1.9, Equation 5.1.10, and Equation 5.1.11. The crosses give the corresponding Fourier coefficients given by the finite Fourier series with $N = 15$.

numerically computed Fourier coefficients in Figure 5.7.2. This figure was created using the MATLAB script

```
clear;
N = 15, M = 2*N; dt = 2*pi/M; % number of points in interval
% create time points assuming x(t) = x(t+period)
t = [-pi:dt:pi-dt];
%
f = zeros(size(t)); % initialize function f(t)
for k = 1:length(t) % construct function f(t)
  if t(k) < 0; f(k) = 0; else f(k) = t(k); end; end;
%
% compute Fourier coefficients using fast Fourier transform
%
fourier = fft(f) / N;
a_0_comp = real(fourier(1)); sign = 1;
for n = 2:N;
  a_n_comp(n-1) = - sign * real(fourier(n));
  b_n_comp(n-1) = sign * imag(fourier(n));
  sign = - sign;
end
%
% plot comparisons
%
NN = linspace(0,N-1,N);
exact_coeff(1) = pi/2;
numer_coeff(1) = a_0_comp;
```

```
for n = 1:N-1;
  exact_coeff(n+1) = ((-1)^n-1) / (pi*(2*n-1)^2);
  numer_coeff(n+1) = a_n_comp(n);
end;
subplot(2,1,1),plot(NN,exact_coeff,'o',NN,numer_coeff,'kx')
ylabel('a_n','Fontsize',20)
clear exact_coeff numer_coeff
NN = linspace(1,N-1,N-1);
for n = 1:N-1;
  exact_coeff(n) = -(-1)^n/n; numer_coeff(n) = b_n_comp(n);
end;
subplot(2,1,2), plot(NN,exact_coeff,'o',NN,numer_coeff,'kx')
xlabel('n','Fontsize',20); ylabel('b_n','Fontsize',20);
```

It shows that a relative few data points can yield quite reasonable answers.

Let us examine this script a little closer. One of the first things that you will note is that there is no explicit reference to Equation 5.7.11 and Equation 5.7.12. How did we get the correct answer?

Although we could have coded Equation 5.7.11 and Equation 5.7.12, no one does that any more. In the 1960s, J. W. Cooley and J. W. Tukey[22] devised an incredibly clever method of performing these calculations. This method, commonly called a fast Fourier transform or FFT, is so popular that all computational packages contain it as an intrinsic function and MATLAB is no exception, calling it `fft`. This is what has been used here.

Although we now have an `fft` to compute the coefficients, this routine does not directly give the coefficients a_n and b_n but rather some mysterious (complex) number that is related to $a_n + ib_n$. This is a common problem in using a package's FFT rather than your own and why the script divides by N and we keep changing the sign. The best method for discovering how to extract the coefficients a_n and b_n is to test it with a dataset created by a simple, finite series such as

$$f(x) = 20 + \cos(t) + 3\sin(t) + 6\cos(2t) - 20\sin(2t) - 10\cos(3t) - 30\sin(3t). \quad \textbf{(5.7.17)}$$

If the code is correct, it must give back the coefficient in Equation 5.7.17 to within round-off. Otherwise, something is wrong.

Finally, most FFTs assume that the dataset will start repeating after the final data point. Therefore, when reading in the dataset, the point corresponding to $x = L$ must be excluded. □

• Example 5.7.3: Aliasing

In the previous example, we could only resolve phenomena with a period of 2 months or greater although we had data for each of the 12 months. This is an example of *Nyquist's sampling criteria:*[23] At least two samples are required to resolve the highest frequency in a periodically sampled record.

Figure 5.7.3 will help explain this phenomenon. In case (a) we have quite a few data points over one cycle. Consequently our picture, constructed from data, is fairly good. In

[22] Cooley, J. W., and J. W. Tukey, 1965: An algorithm for machine calculation of complex Fourier series. *Math. Comput.*, **19**, 297–301.

[23] Nyquist, H., 1928: Certain topics in telegraph transmission theory. *AIEE Trans.*, **47**, 617–644.

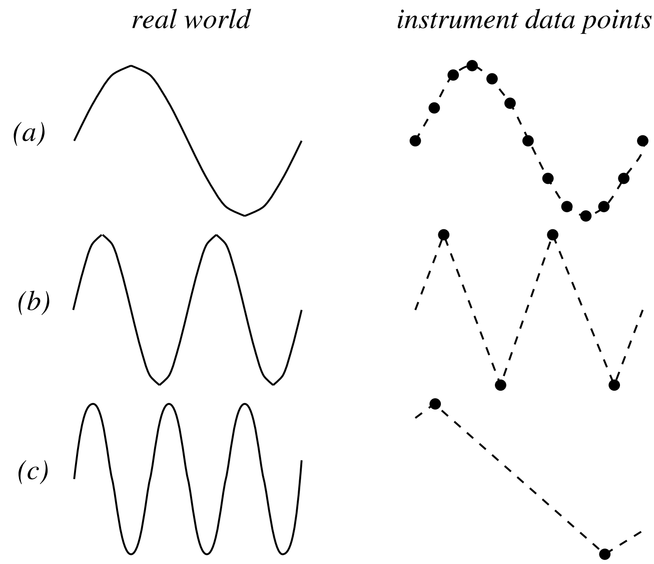


Figure 5.7.3: The effect of sampling in the representation of periodic functions.

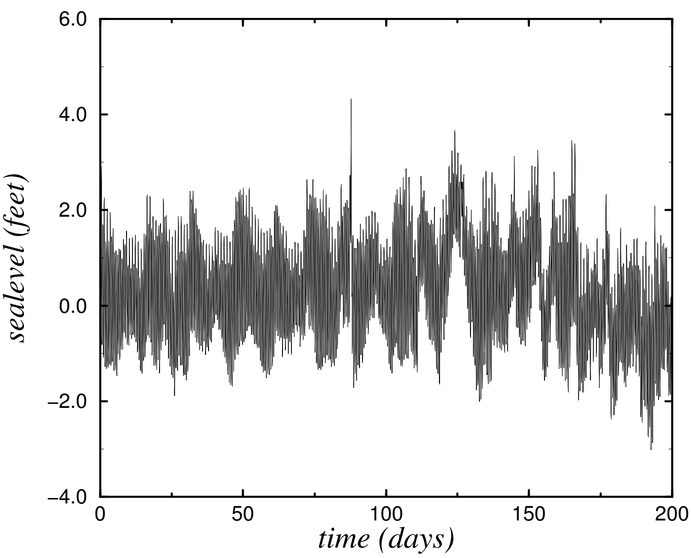


Figure 5.7.4: The sea elevation at the mouth of the Chesapeake Bay from its average depth as a function of time after 1 July 1985.

case (b), we took only samples at the ridges and troughs of the wave. Although our picture of the real phenomenon is poor, at least we know that there is a wave. From this picture we see that even if we are lucky enough to take our observations at the ridges and troughs of a wave, we need at least two data points per cycle (one for the ridge, the other for the trough) to resolve the highest-frequency wave.

In case (c) we have made a big mistake. We have taken a wave of frequency N Hz and misrepresented it as a wave of frequency $N/2$ Hz. This misrepresentation of a high-frequency wave by a lower-frequency wave is called *aliasing*. It arises because we are sampling a continuous signal at equal intervals. By comparing cases (b) and (c), we see that there is

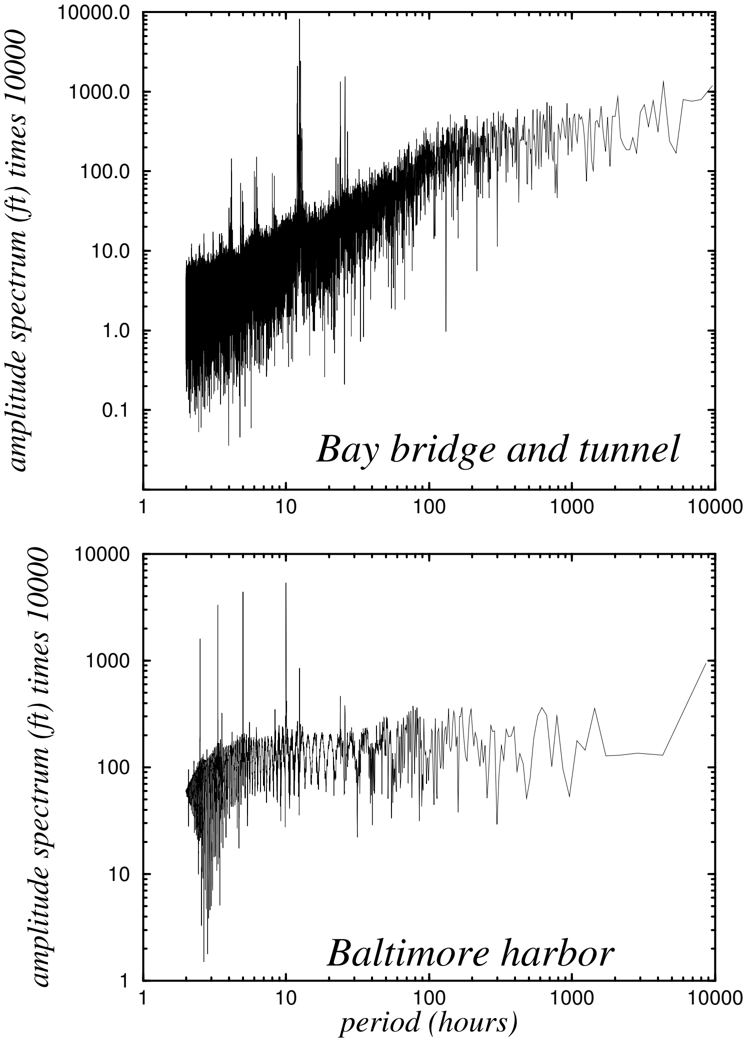


Figure 5.7.5: The amplitude of the Fourier coefficients for the sea elevation at the Chesapeake Bay bridge and tunnel (top) and Baltimore harbor (bottom) as a function of period.

a cutoff between aliased and nonaliased frequencies. This frequency is called the *Nyquist* or *folding* frequency. It corresponds to the highest frequency resolved by our finite Fourier analysis.

Because most periodic functions require an infinite number of harmonics for their representation, aliasing of signals is a common problem. Thus the question is not "can I avoid aliasing?" but "can I live with it?" Quite often, we can construct our experiments to say yes. An example where aliasing is unavoidable occurs in a Western at the movies when we see the rapidly rotating spokes of the stagecoach's wheel. A movie is a sampling of continuous motion where we present the data as a succession of pictures. Consequently, a film aliases the high rate of revolution of the stagecoach's wheel in such a manner so that it appears to be stationary or rotating very slowly. □

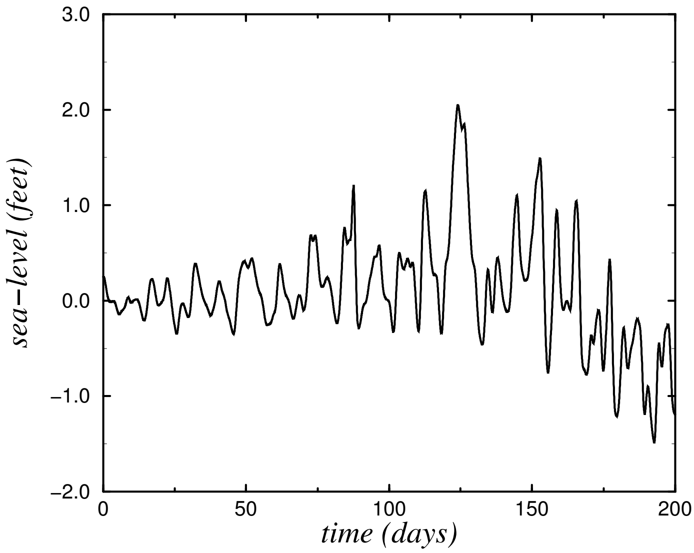


Figure 5.7.6: Same as Figure 5.7.4 but with the tides removed.

• **Example 5.7.4: Spectrum of the Chesapeake Bay**

For our final example, we perform a Fourier analysis of hourly sea-level measurements taken at the mouth of the Chesapeake Bay during the 2000 days from 9 April 1985 to 29 June 1990. Figure 5.7.4 shows 200 days of this record, starting from 1 July 1985. As this figure shows, the measurements contain a wide range of oscillations. In particular, note the large peak near day 90 that corresponds to the passage of Hurricane Gloria during the early hours of 27 September 1985.

Utilizing the entire 2000 days, we plotted the amplitude of the Fourier coefficients as a function of period in Figure 5.7.5. We see a general rise of the amplitude as the period increases. Especially noteworthy are the sharp peaks near periods of 12 and 24 hours. The largest peak is at 12.417 hours and corresponds to the semidiurnal tide. Thus, our Fourier analysis shows that the dominant oscillations at the mouth of the Chesapeake Bay are the tides. A similar situation occurs in Baltimore harbor. Furthermore, with this spectral information we could predict high and low tides very accurately.

Although the tides are of great interest to some, they are a nuisance to others because they mask other physical processes that might be occurring. For that reason we would like to remove them from the tidal gauge history and see what is left. One way would be to zero out the Fourier coefficients corresponding to the tidal components and then plot the resulting Fourier series. Another method is to replace each hourly report with an average of hourly reports that occurred 24 hours ahead of and behind a particular report. We construct this average in such a manner that waves with periods of the tides sum to zero.[24] Such a *filter* is a popular method for eliminating unwanted waves from a record. Filters play an important role in the analysis of data. We plotted the filtered sea level data in Figure 5.7.6. Note that summertime (0–50 days) produces little variation in the sea level compared to wintertime (100–150 days) when intense coastal storms occur.

---

[24] See Godin, G., 1972: *The Analysis of Tides*. University of Toronto Press, Section 2.1.

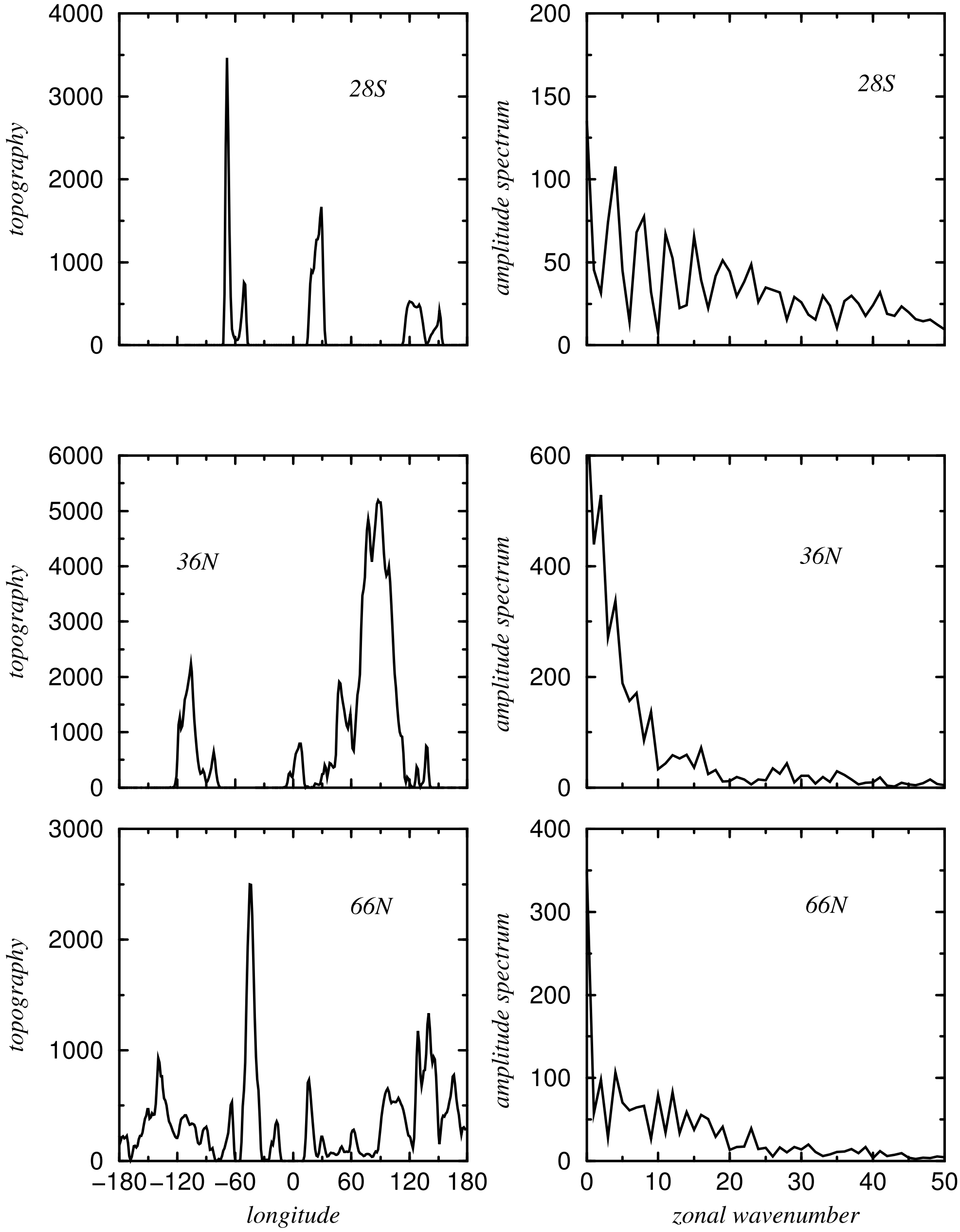


Figure 5.7.7: The orography of the earth and its spectrum in meters along three latitude belts using a topography dataset with a resolution of 1.25° longitude.

Problems

Find the finite Fourier series for the following pieces of data:

1. $f(0) = 0$, $f(1) = 1$, $f(2) = 2$, $f(3) = 3$, and $N = 2$.

2. $f(0) = 1$, $f(1) = 1$, $f(2) = -1$, $f(3) = -1$, and $N = 2$.

Project: Spectrum of the Earth's Orography

Table 5.7.2 gives the orographic height of the earth's surface used in an atmospheric general circulation model (GCM) at a resolution of 2.5° longitude along the latitude belts of 28°S, 36°N, and 66°N. In this project you will find the spectrum of this orographic field along the various latitude belts.

Step 1: Write a MATLAB script that reads in the data and find A_n and B_n and then construct the amplitude spectra for this data.

Step 2: Construct several spectra by using every data point, every other data point, etc. How do the magnitudes of the Fourier coefficient change? You might like to read about *leakage* from a book on harmonic analysis.[25]

Step 3: Compare and contrast the spectra from the various latitude belts. How do the magnitudes of the Fourier coefficients decrease with n? Why are there these differences?

Step 4: You may have noted that some of the heights are negative, even in the middle of the ocean! Take the original data (for any latitude belt) and zero out all of the negative heights. Find the spectra for this new data set. How have the spectra changed? Is there a reason why the negative heights were introduced?

Further Readings

Carslaw, H. S., 1950: *An Introduction to the Theory of Fourier's Series and Integrals.* Dover, 368 pp. A classic treatment of the Fourier technique.

Tolstov, Georgi P., 1976: *Fourier Series.* Dover, 336 pp. This book covers the basic theory of Fourier series and its use in mathematical physics.

[25] For example, Bloomfield, P., 1976: *Fourier Analysis of Time Series: An Introduction*. John Wiley & Sons, 258 pp.

Table 5.7.2: Orographic Heights (in m) Times the Gravitational Acceleration Constant ($g = 9.81$ m/s^2) along Three Latitude Belts

Longitude	28°S	36°N	66°N	Longitude	28°S	36°N	66°N
−180.0	4.	3.	2532.	−82.5	36.	4047.	737.
−177.5	1.	−2.	1665.	−80.0	−64.	3938.	185.
−175.0	1.	2.	1432.	−77.5	138.	1669.	71.
−172.5	1.	−3.	1213.	−75.0	−363.	236.	160.
−170.0	1.	1.	501.	−72.5	4692.	31.	823.
−167.5	1.	−3.	367.	−70.0	19317.	−8.	1830.
−165.0	1.	1.	963.	−67.5	21681.	0.	3000.
−162.5	0.	0.	1814.	−65.0	9222.	−2.	3668.
−160.0	−1.	6.	2562.	−62.5	1949.	−2.	2147.
−157.5	0.	1.	3150.	−60.0	774.	0.	391.
−155.0	0.	3.	4008.	−57.5	955.	5.	−77.
−152.5	1.	−2.	4980.	−55.0	2268.	6.	601.
−150.0	−1.	4.	6011.	−52.5	4636.	−1.	3266.
−147.5	6.	−1.	6273.	−50.0	4621.	2.	9128.
−145.0	14.	3.	5928.	−47.5	1300.	−4.	17808.
−142.5	6.	−1.	6509.	−45.0	−91.	1.	22960.
−140.0	−2.	6.	7865.	−42.5	57.	−1.	20559.
−137.5	0.	3.	7752.	−40.0	−25.	4.	14296.
−135.0	−2.	5.	6817.	−37.5	13.	−1.	9783.
−132.5	1.	−2.	6272.	−35.0	−10.	6.	5969.
−130.0	−2.	0.	5582.	−32.5	8.	2.	1972.
−127.5	0.	5.	4412.	−30.0	−4.	22.	640.
−125.0	−2.	423.	3206.	−27.5	6.	33.	379.
−122.5	1.	3688.	2653.	−25.0	−2.	39.	286.
−120.0	−3.	10919.	2702.	−22.5	3.	2.	981.
−117.5	2.	16148.	3062.	−20.0	−3.	11.	1971.
−115.0	−3.	17624.	3344.	−17.5	1.	−6.	2576.
−112.5	7.	18132.	3444.	−15.0	−1.	19.	1692.
−110.0	12.	19511.	3262.	−12.5	0.	−18.	357.
−107.5	9.	22619.	3001.	−10.0	−1.	490.	−21.
−105.0	−5.	20273.	2931.	−7.5	0.	2164.	−5.
−102.5	3.	12914.	2633.	−5.0	1.	4728.	−10.
−100.0	−5.	7434.	1933.	−2.5	0.	5347.	0.
−97.5	6.	4311.	1473.	0.0	4.	2667.	−6.
−95.0	−8.	2933.	1689.	2.5	−5.	1213.	−1.
−92.5	8.	2404.	2318.	5.0	7.	1612.	−31.
−90.0	−12.	1721.	2285.	7.5	−13.	1744.	−58.
−87.5	18.	1681.	1561.	10.0	28.	1153.	381.
−85.0	−23.	2666.	1199.	12.5	107.	838.	2472.
15.0	2208.	1313.	5263.	97.5	0.	35538.	6222.
17.5	6566.	862.	5646.	100.0	−2.	31985.	5523.
20.0	9091.	1509.	3672.	102.5	0.	23246.	4823.
22.5	10690.	2483.	1628.	105.0	−4.	17363.	4689.
25.0	12715.	1697.	889.	107.5	2.	14315.	4698.

Table 5.7.2, contd.: Orographic Heights (in m) Times the Gravitational Acceleration Constant ($g = 9.81$ m/s^2) along Three Latitude Belts

Longitude	28°S	36°N	66°N	Longitude	28°S	36°N	66°N
27.5	14583.	3377.	1366.	110.0	−17.	12639.	4674.
30.0	11351.	7682.	1857.	112.5	302.	10543.	4435.
32.5	3370.	9663.	1534.	115.0	1874.	4967.	3646.
35.0	15.	10197.	993.	117.5	4005.	1119.	2655.
37.5	49.	10792.	863.	120.0	4989.	696.	2065.
40.0	−31.	11322.	756.	122.5	4887.	475.	1583.
42.5	20.	13321.	620.	125.0	4445.	1631.	3072.
45.0	−17.	15414.	626.	127.5	4362.	2933.	7290.
47.5	−19.	12873.	836.	130.0	4368.	1329.	8541.
50.0	−18.	6114.	1029.	132.5	3485.	88.	7078.
52.5	6.	2962.	946.	135.0	1921.	598.	7322.
55.0	−2.	4913.	828.	137.5	670.	1983.	9445.
57.5	3.	6600.	1247.	140.0	666.	2511.	10692.
60.0	−3.	4885.	2091.	142.5	1275.	866.	9280.
62.5	2.	3380.	2276.	145.0	1865.	13.	8372.
65.0	−1.	5842.	1870.	147.5	2452.	11.	6624.
67.5	2.	12106.	1215.	150.0	3160.	−4.	3617.
70.0	0.	23032.	680.	152.5	2676.	−1.	2717.
72.5	2.	35376.	531.	155.0	697.	0.	3474.
75.0	−1.	36415.	539.	157.5	−67.	−3.	4337.
77.5	1.	26544.	579.	160.0	25.	3.	4824.
80.0	0.	19363.	554.	162.5	−12.	−1.	5525.
82.5	1.	17915.	632.	165.0	10.	4.	6323.
85.0	−2.	22260.	791.	167.5	−5.	−2.	5899.
87.5	−1.	30442.	1455.	170.0	0.	1.	4330.
90.0	−3.	33601.	3194.	172.5	0.	−4.	3338.
92.5	−1.	30873.	4878.	175.0	4.	3.	3408.
95.0	0.	31865.	5903.	177.5	3.	−1.	3407.

Chapter 6
The Sturm-Liouville Problem

In the next three chapters we will be solving partial differential equations using the technique of separation of variables. This technique requires that we expand a piece-wise continuous function $f(x)$ as a linear sum of *eigenfunctions*, much as we used sines and cosines to re-express $f(x)$ in a Fourier series. The purpose of this chapter is to explain and illustrate these eigenfunction expansions.

6.1 EIGENVALUES AND EIGENFUNCTIONS

Repeatedly, in the next three chapters on partial differential equations, we will solve the following second-order linear differential equation:

$$\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + [q(x) + \lambda r(x)]y = 0, \quad a \le x \le b, \tag{6.1.1}$$

together with the boundary conditions:

$$\alpha y(a) + \beta y'(a) = 0 \qquad \text{and} \qquad \gamma y(b) + \delta y'(b) = 0. \tag{6.1.2}$$

In Equation 6.1.1, $p(x)$, $q(x)$, and $r(x)$ are real functions of x; λ is a parameter; and $p(x)$ and $r(x)$ are functions that are continuous and positive on the interval $a \le x \le b$.

By the time that Charles-François Sturm (1803–1855) met Joseph Liouville in the early 1830s, he had already gained fame for his work on the compression of fluids and his celebrated theorem on the number of real roots of a polynomial. An eminent teacher, Sturm spent most of his career teaching at various Parisian colleges. (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

Taken together, Equation 6.1.1 and Equation 6.1.2 constitute a regular *Sturm-Liouville problem*, named after the French mathematicians Sturm and Liouville[1] who first studied these equations in the 1830s. In the case when $p(x)$ or $r(x)$ vanishes at one of the endpoints of the interval $[a, b]$ or when the interval is of infinite length, the problem becomes a *singular Sturm-Liouville problem*.

Consider now the solutions to the regular Sturm-Liouville problem. Clearly there is the trivial solution $y = 0$ for all λ. However, nontrivial solutions exist only if λ takes on specific values; these values are called *characteristic values* or *eigenvalues*. The corresponding nontrivial solutions are called the *characteristic! functions* or *eigenfunctions*. In particular, we have the following theorems.

Theorem: *For a regular Sturm-Liouville problem with $p(x) > 0$, all of the eigenvalues are real if $p(x)$, $q(x)$, and $r(x)$ are real functions and the eigenfunctions are differentiable and continuous.*

Proof: Let $y(x) = u(x) + iv(x)$ be an eigenfunction corresponding to an eigenvalue $\lambda =$

[1] For the complete history as well as the relevant papers, see Lützen, J., 1984: Sturm and Liouville's work on ordinary linear differential equations. The emergence of Sturm-Liouville theory. *Arch. Hist. Exact Sci.*, **29**, 309–376.

Although educated as an engineer, Joseph Liouville (1809–1882) would devote his life to teaching pure and applied mathematics in the leading Parisian institutions of higher education. Today he is most famous for founding and editing for almost 40 years the *Journal de Liouville.* (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

$\lambda_r + i\lambda_i$, where λ_r, λ_i are real numbers and $u(x), v(x)$ are real functions of x. Substituting into the Sturm-Liouville equation yields

$$\{p(x)[u'(x) + iv'(x)]\}' + [q(x) + (\lambda_r + i\lambda_i)r(x)][u(x) + iv(x)] = 0. \qquad \textbf{(6.1.3)}$$

Separating the real and imaginary parts gives

$$[p(x)u'(x)]' + [q(x) + \lambda_r]u(x) - \lambda_i r(x)v(x) = 0, \qquad \textbf{(6.1.4)}$$

and

$$[p(x)v'(x)]' + [q(x) + \lambda_r]v(x) + \lambda_i r(x)u(x) = 0. \qquad \textbf{(6.1.5)}$$

If we multiply Equation 6.1.4 by v and Equation 6.1.5 by u and subtract the results, we find that

$$u(x)[p(x)v'(x)]' - v(x)[p(x)u'(x)]' + \lambda_i r(x)[u^2(x) + v^2(x)] = 0. \qquad \textbf{(6.1.6)}$$

The derivative terms in Equation 6.1.6 can be rewritten so that it becomes

$$\frac{d}{dx}\{[p(x)v'(x)]u(x) - [p(x)u'(x)]v(x)\} + \lambda_i r(x)[u^2(x) + v^2(x)] = 0. \qquad \textbf{(6.1.7)}$$

Integrating from a to b, we find that

$$-\lambda_i \int_a^b r(x)[u^2(x) + v^2(x)]\,dx = \{p(x)[u(x)v'(x) - v(x)u'(x)]\}\big|_a^b. \qquad \textbf{(6.1.8)}$$

From the boundary conditions, Equation 6.1.2,

$$\alpha[u(a)+iv(a)]+\beta[u'(a)+iv'(a)]=0, \tag{6.1.9}$$

and

$$\gamma[u(b)+iv(b)]+\delta[u'(b)+iv'(b)]=0. \tag{6.1.10}$$

Separating the real and imaginary parts yields

$$\alpha u(a)+\beta u'(a)=0, \quad \text{and} \quad \alpha v(a)+\beta v'(a)=0, \tag{6.1.11}$$

and

$$\gamma u(b)+\delta u'(b)=0, \quad \text{and} \quad \gamma v(b)+\delta v'(b)=0. \tag{6.1.12}$$

Both α and β cannot be zero; otherwise, there would be no boundary condition at $x=a$. Similar considerations hold for γ and δ. Therefore,

$$u(a)v'(a)-u'(a)v(a)=0, \quad \text{and} \quad u(b)v'(b)-u'(b)v(b)=0, \tag{6.1.13}$$

if we treat α, β, γ, and δ as unknowns in a system of homogeneous equations, Equation 6.1.11 and Equation 6.1.12, and require that the corresponding determinants equal zero. Applying Equation 6.1.13 to the right side of Equation 6.1.8, we obtain

$$\lambda_i \int_a^b r(x)[u^2(x)+v^2(x)]\,dx=0. \tag{6.1.14}$$

Because $r(x)>0$, the integral is positive and $\lambda_i=0$. Since $\lambda_i=0$, λ is purely real. This implies that the eigenvalues are real. □

If there is only one independent eigenfunction for each eigenvalue, that eigenvalue is *simple*. When more than one eigenfunction belongs to a single eigenvalue, the problem is *degenerate*.

Theorem: *The regular Sturm-Liouville problem has infinitely many real and simple eigenvalues* λ_n, $n=0,1,2,\ldots$, *which can be arranged in a monotonically increasing sequence* $\lambda_0<\lambda_1<\lambda_2<\cdots$ *such that* $\lim_{n\to\infty}\lambda_n=\infty$. *Every eigenfunction* $y_n(x)$ *associated with the corresponding eigenvalue* λ_n *has exactly* n *zeros in the interval* (a,b). *For each eigenvalue there exists only one eigenfunction (up to a multiplicative constant).*

The proof is beyond the scope of this book but may be found in more advanced treatises.[2]
□

In the following examples we illustrate how to find these real eigenvalues and their corresponding eigenfunctions.

• Example 6.1.1

Let us find the eigenvalues and eigenfunctions of

$$y''+\lambda y=0, \tag{6.1.15}$$

[2] See, for example, Birkhoff, G., and G.-C. Rota, 1989: *Ordinary Differential Equations*. John Wiley & Sons, Chapters 10 and 11; Sagan, H., 1961: *Boundary and Eigenvalue Problems in Mathematical Physics*. John Wiley & Sons, Chapter 5.

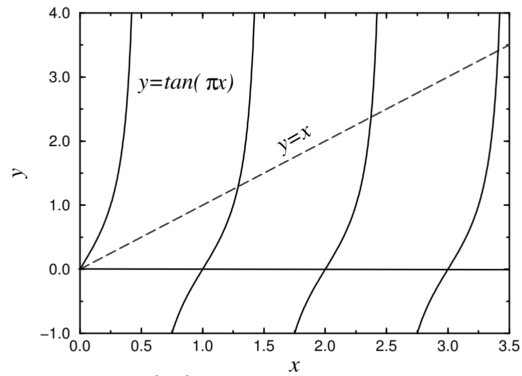


Figure 6.1.1: Graphical solution of $\tan(\pi x) = x$.

subject to the boundary conditions

$$y(0) = 0, \qquad \text{and} \qquad y(\pi) - y'(\pi) = 0. \tag{6.1.16}$$

Our first task is to check to see whether the problem is indeed a regular Sturm-Liouville problem. A comparison between Equation 6.1.1 and Equation 6.1.15 shows that they are the same if $p(x) = 1$, $q(x) = 0$, and $r(x) = 1$. Similarly, the boundary conditions, Equation 6.1.16, are identical to Equation 6.1.2 if $\alpha = \gamma = 1$, $\delta = -1$, $\beta = 0$, $a = 0$, and $b = \pi$.

Because the form of the solution to Equation 6.1.15 depends on λ, we consider three cases: λ negative, positive, or equal to zero. The general solution[3] of the differential equation is

$$y(x) = A\cosh(mx) + B\sinh(mx), \quad \text{if} \quad \lambda < 0, \tag{6.1.17}$$

$$y(x) = C + Dx, \quad \text{if} \quad \lambda = 0, \tag{6.1.18}$$

and

$$y(x) = E\cos(kx) + F\sin(kx), \quad \text{if} \quad \lambda > 0, \tag{6.1.19}$$

where for convenience $\lambda = -m^2 < 0$ in Equation 6.1.17 and $\lambda = k^2 > 0$ in Equation 6.1.19. Both k and m are real and positive by these definitions.

Turning to the condition that $y(0) = 0$, we find that $A = C = E = 0$. The other boundary condition $y(\pi) - y(\pi) = 0$ gives

$$B[\sinh(m\pi) - m\cosh(m\pi)] = 0, \tag{6.1.20}$$

$$D = 0, \tag{6.1.21}$$

and

$$F[\sin(k\pi) - k\cos(k\pi)] = 0. \tag{6.1.22}$$

If we graph $\sinh(m\pi) - m\cosh(m\pi)$ for all positive m, this quantity is always negative. Consequently, $B = 0$. However, in Equation 6.1.22, a nontrivial solution (i.e., $F \neq 0$) occurs if

$$F\cos(k\pi)[\tan(k\pi) - k] = 0, \quad \text{or} \quad \tan(k\pi) = k. \tag{6.1.23}$$

[3] In many differential equations courses, the solution to $y'' - m^2y = 0$, $m > 0$ is written $y(x) = c_1e^{mx} + c_2e^{-mx}$. However, we can rewrite this solution as $y(x) = (c_1+c_2)\frac{1}{2}(e^{mx}+e^{-mx})+(c_1-c_2)\frac{1}{2}(e^{mx}-e^{-mx}) = A\cosh(mx) + B\sinh(mx)$, where $\cosh(mx) = (e^{mx} + e^{-mx})/2$ and $\sinh(mx) = (e^{mx} - e^{-mx})/2$. The advantage of using these hyperbolic functions over exponentials is the simplification that occurs when we substitute the hyperbolic functions into the boundary conditions.

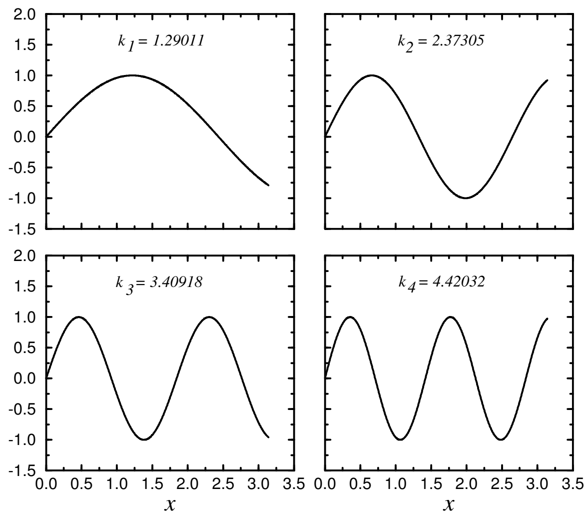


Figure 6.1.2: The first four eigenfunctions $\sin(k_n x)$ corresponding to the eigenvalue problem $\tan(k\pi) = k$.

In summary, we found nontrivial solutions only when $\lambda_n = k_n^2 > 0$, where k_n is the nth root of the transcendental equation, Equation 6.1.23. We can find the roots either graphically or through the use of a numerical algorithm. Figure 6.1.1 illustrates the graphical solution to the problem. We exclude the root $k = 0$ because λ must be greater than zero.

Let us now find the corresponding eigenfunctions. Because $A = B = C = D = E = 0$, we are left with $y(x) = F\sin(kx)$. Consequently, the eigenfunction, traditionally written without the arbitrary amplitude constant, is

$$y_n(x) = \sin(k_n x), \tag{6.1.24}$$

because k must equal k_n. Figure 6.1.2 shows the first four eigenfunctions. □

• Example 6.1.2

For our second example let us solve the Sturm-Liouville problem,[4]

$$y'' + \lambda y = 0, \tag{6.1.25}$$

with the boundary conditions

$$y(0) - y'(0) = 0, \quad \text{and} \quad y(\pi) - y'(\pi) = 0. \tag{6.1.26}$$

Once again the three possible solutions to Equation 6.1.25 are

$$y(x) = A\cosh(mx) + B\sinh(mx), \quad \text{if} \quad \lambda = -m^2 < 0, \tag{6.1.27}$$

$$y(x) = C + Dx, \quad \text{if} \quad \lambda = 0, \tag{6.1.28}$$

[4] Sosov and Theodosiou [Sosov, Y., and C. E. Theodosiou, 2002: On the complete solution of the Sturm-Liouville problem $(d^2X/dx^2) + \lambda^2 X = 0$ over a closed interval. *J. Math. Phys. (Woodbury, NY)*, **43**, 2831–2843] have analyzed this problem with the general boundary conditions, Equation 6.1.2.

and

$$y(x) = E\cos(kx) + F\sin(kx), \quad \text{if} \quad \lambda = k^2 > 0. \tag{6.1.29}$$

Let us first check and see if there are any nontrivial solutions for $\lambda < 0$. Two simultaneous equations result from the substitution of Equation 6.1.27 into Equation 6.1.26:

$$A - mB = 0, \tag{6.1.30}$$

and

$$[\cosh(m\pi) - m\sinh(m\pi)]A + [\sinh(m\pi) - m\cosh(m\pi)]B = 0. \tag{6.1.31}$$

The elimination of A between the two equations yields

$$\sinh(m\pi)(1 - m^2)B = 0. \tag{6.1.32}$$

If Equation 6.1.27 is a nontrivial solution, then $B \neq 0$, and

$$\sinh(m\pi) = 0, \qquad \text{or} \qquad m^2 = 1. \tag{6.1.33}$$

The condition $\sinh(m\pi) = 0$ cannot hold because it implies $m = \lambda = 0$, which contradicts the assumption used in deriving Equation 6.1.27 that $\lambda < 0$. On the other hand, $m^2 = 1$ is quite acceptable. It corresponds to the eigenvalue $\lambda = -1$ and the eigenfunction is

$$y_0 = \cosh(x) + \sinh(x) = e^x, \tag{6.1.34}$$

because it satisfies the differential equation

$$y_0'' - y_0 = 0, \tag{6.1.35}$$

and the boundary conditions

$$y_0(0) - y_0'(0) = 0, \qquad \text{and} \qquad y_0(\pi) - y_0'(\pi) = 0. \tag{6.1.36}$$

An alternative method of finding m, which is quite popular because of its use in more difficult problems, follows from viewing Equation 6.1.30 and Equation 6.1.31 as a system of homogeneous linear equations, where A and B are the unknowns. It is well known[5] that for Equation 6.1.30 and Equation 6.1.31 to have a nontrivial solution (i.e., $A \neq 0$ and/or $B \neq 0$) the determinant of the coefficients must vanish:

$$\begin{vmatrix} 1 & -m \\ \cosh(m\pi) - m\sinh(m\pi) & \sinh(m\pi) - m\cosh(m\pi) \end{vmatrix} = 0. \tag{6.1.37}$$

Expanding the determinant,

$$\sinh(m\pi)(1 - m^2) = 0, \tag{6.1.38}$$

which leads directly to Equation 6.1.33.

We consider next the case of $\lambda = 0$. Substituting Equation 6.1.28 into Equation 6.1.26, we find that

$$C - D = 0, \qquad \text{and} \qquad C + D\pi - D = 0. \tag{6.1.39}$$

---

[5] See Chapter 3.

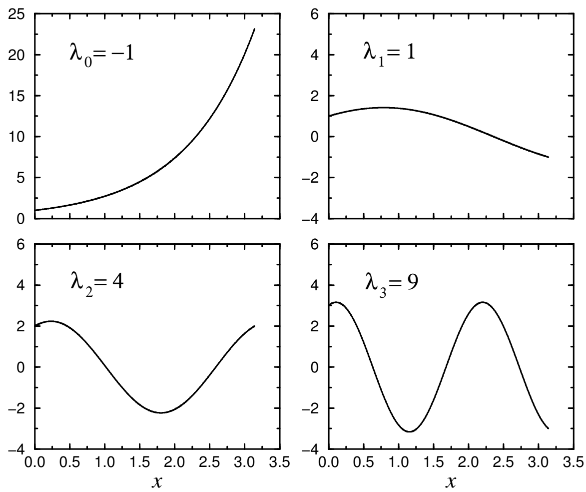


Figure 6.1.3: The first four eigenfunctions for the Sturm-Liouville problem, Equation 6.1.25 and Equation 6.1.26.

This set of simultaneous equations yields $C = D = 0$ and we have only trivial solutions for $\lambda = 0$.

Finally, we examine the case when $\lambda > 0$. Substituting Equation 6.1.29 into Equation 6.1.26, we obtain

$$E - kF = 0, \tag{6.1.40}$$

and

$$[\cos(k\pi) + k\sin(k\pi)]E + [\sin(k\pi) - k\cos(k\pi)]F = 0. \tag{6.1.41}$$

The elimination of E from Equation 6.1.40 and Equation 6.1.41 gives

$$F(1 + k^2)\sin(k\pi) = 0. \tag{6.1.42}$$

If Equation 6.1.29 is nontrivial, $F \neq 0$, and

$$k^2 = -1, \qquad \text{or} \qquad \sin(k\pi) = 0. \tag{6.1.43}$$

The condition $k^2 = -1$ violates the assumption that k is real, which follows from the fact that $\lambda = k^2 > 0$. On the other hand, we can satisfy $\sin(k\pi) = 0$ if $k = 1, 2, 3, \ldots$; a negative k yields the same λ. Consequently we have the additional eigenvalues $\lambda_n = n^2$.

Let us now find the corresponding eigenfunctions. Because $E = kF$, $y(x) = F\sin(kx) + Fk\cos(kx)$ from Equation 6.1.29. Thus, the eigenfunctions for $\lambda > 0$ are

$$y_n(x) = \sin(nx) + n\cos(nx). \tag{6.1.44}$$

Figure 6.1.3 illustrates some of the eigenfunctions given by Equation 6.1.34 and Equation 6.1.44. □

• **Example 6.1.3**

Consider now the Sturm-Liouville problem

$$y'' + \lambda y = 0, \tag{6.1.45}$$

with

$$y(\pi) = y(-\pi), \quad \text{and} \quad y'(\pi) = y'(-\pi). \tag{6.1.46}$$

This is *not* a regular Sturm-Liouville problem because the boundary conditions are periodic and do not conform to the canonical boundary condition, Equation 6.1.2.

The general solution to Equation 6.1.45 is

$$y(x) = A\cosh(mx) + B\sinh(mx), \quad \text{if} \quad \lambda = -m^2 < 0, \tag{6.1.47}$$

$$y(x) = C + Dx, \quad \text{if} \quad \lambda = 0, \tag{6.1.48}$$

and

$$y(x) = E\cos(kx) + F\sin(kx), \quad \text{if} \quad \lambda = k^2 > 0. \tag{6.1.49}$$

Substituting these solutions into the boundary condition, Equation 6.1.46,

$$A\cosh(m\pi) + B\sinh(m\pi) = A\cosh(-m\pi) + B\sinh(-m\pi), \tag{6.1.50}$$

$$C + D\pi = C - D\pi, \tag{6.1.51}$$

and

$$E\cos(k\pi) + F\sin(k\pi) = E\cos(-k\pi) + F\sin(-k\pi), \tag{6.1.52}$$

or

$$B\sinh(m\pi) = 0, \quad D = 0, \quad \text{and} \quad F\sin(k\pi) = 0, \tag{6.1.53}$$

because $\cosh(-m\pi) = \cosh(m\pi)$, $\sinh(-m\pi) = -\sinh(m\pi)$, $\cos(-k\pi) = \cos(k\pi)$, and $\sin(-k\pi) = -\sin(k\pi)$. Because m must be positive, $\sinh(m\pi)$ cannot equal zero and $B = 0$. On the other hand, if $\sin(k\pi) = 0$ or $k = n$, $n = 1, 2, 3, \ldots$, we have a nontrivial solution for positive λ and $\lambda_n = n^2$. Note that we still have A, C, E, and F as free constants.

From the boundary condition, Equation 6.1.46,

$$A\sinh(m\pi) = A\sinh(-m\pi), \tag{6.1.54}$$

and

$$-E\sin(k\pi) + F\cos(k\pi) = -E\sin(-k\pi) + F\cos(-k\pi). \tag{6.1.55}$$

The solution $y_0(x) = C$ identically satisfies the boundary condition, Equation 6.1.46, for all C. Because m and $\sinh(m\pi)$ must be positive, $A = 0$. From Equation 6.1.53, we once again have $\sin(k\pi) = 0$, and $k = n$. Consequently, the eigenfunction solutions to Equation 6.1.45 and Equation 6.1.46 are

$$\lambda_0 = 0, \qquad y_0(x) = 1, \tag{6.1.56}$$

and

$$\lambda_n = n^2, \qquad y_n(x) = \begin{cases} \sin(nx), \\ \cos(nx), \end{cases} \tag{6.1.57}$$

and we have a degenerate set of eigenfunctions to the Sturm-Liouville problem, Equation 6.1.45, with the periodic boundary condition, Equation 6.1.46.

Problems

Find the eigenvalues and eigenfunctions for each of the following:

1. $y'' + \lambda y = 0, \quad y'(0) = 0, \quad y(L) = 0$

2. $y'' + \lambda y = 0, \quad y'(0) = 0, \quad y'(\pi) = 0$

3. $y'' + \lambda y = 0, \quad y(0) + y'(0) = 0, \quad y(\pi) + y'(\pi) = 0$

4. $y'' + \lambda y = 0, \quad y'(0) = 0, \quad y(\pi) - y'(\pi) = 0$

5. $y^{(iv)} + \lambda y = 0, \quad y(0) = y''(0) = 0, \quad y(L) = y''(L) = 0$

Find an equation from which you could find λ and give the form of the eigenfunction for each of the following:

6. $y'' + \lambda y = 0, \quad y(0) + y'(0) = 0, \quad y(1) = 0$

7. $y'' + \lambda y = 0, \quad y(0) = 0, \quad y(\pi) + y'(\pi) = 0$

8. $y'' + \lambda y = 0, \quad y'(0) = 0, \quad y(1) - y'(1) = 0$

9. $y'' + \lambda y = 0, \quad y(0) + y'(0) = 0, \quad y'(\pi) = 0$

10. $y'' + \lambda y = 0, \quad y(0) + y'(0) = 0, \quad y(\pi) - y'(\pi) = 0$

11. Find the eigenvalues and eigenfunctions of the Sturm-Liouville problem

$$\frac{d}{dx}\left(x\frac{dy}{dx}\right) + \frac{\lambda}{x}y = 0, \quad 1 \le x \le e$$

for each of the following boundary conditions: (a) $u(1) = u(e) = 0$, (b) $u(1) = u'(e) = 0$, and (c) $u'(1) = u'(e) = 0$.

Find the eigenvalues and eigenfunctions of the following Sturm-Liouville problems:

12.
$$x^2y'' + 2xy' + \lambda y = 0, \quad y(1) = y(e) = 0, \quad 1 \le x \le e$$

13.
$$\frac{d}{dx}\left(x^3y'\right) + \lambda xy = 0, \quad y(1) = y(e^\pi) = 0, \quad 1 \le x \le e^\pi$$

14.
$$\frac{d}{dx}\left(\frac{1}{x}y'\right) + \frac{\lambda}{x}y = 0, \quad y(1) = y(e) = 0, \quad 1 \le x \le e$$

15.
$$y'''' - \lambda^4 y = 0, \quad y'''(0) = y''(0) = y'''(1) = y'(1) = 0, \quad 0 < x < 1$$

6.2 ORTHOGONALITY OF EIGENFUNCTIONS

In the previous section we saw how nontrivial solutions to the regular Sturm-Liouville problem consist of eigenvalues and eigenfunctions. The most important property of eigenfunctions is orthogonality.

Theorem: *Let the functions $p(x)$, $q(x)$, and $r(x)$ of the regular Sturm-Liouville problem, Equation 6.1.1 and Equation 6.1.2, be real and continuous on the interval $[a, b]$. If $y_n(x)$ and $y_m(x)$ are continuously differentiable eigenfunctions corresponding to the distinct eigenvalues λ_n and λ_m, respectively, then $y_n(x)$ and $y_m(x)$ satisfy the* orthogonality *condition:*

$$\boxed{\int_a^b r(x) y_n(x) y_m(x)\, dx = 0,} \qquad \textbf{(6.2.1)}$$

if $\lambda_n \neq \lambda_m$. When Equation 6.2.1 is satisfied, the eigenfunctions $y_n(x)$ and $y_m(x)$ are said to be *orthogonal* to each other with respect to the *weight function* $r(x)$. The term *orthogonality* appears to be borrowed from linear algebra where a similar relationship holds between two perpendicular or orthogonal vectors.

Proof: Let $y_n(x)$ and $y_m(x)$ denote the eigenfunctions associated with two different eigenvalues λ_n and λ_m. Then

$$\frac{d}{dx}\left[p(x)\frac{dy_n}{dx}\right] + [q(x) + \lambda_n r(x)]y_n(x) = 0, \qquad \textbf{(6.2.2)}$$

$$\frac{d}{dx}\left[p(x)\frac{dy_m}{dx}\right] + [q(x) + \lambda_m r(x)]y_m(x) = 0, \qquad \textbf{(6.2.3)}$$

and both solutions satisfy the boundary conditions. Let us multiply the first differential equation by y_m; the second by y_n. Next, we subtract these two equations and move the terms containing $y_n y_m$ to the right side. The resulting equation is

$$y_n \frac{d}{dx}\left[p(x)\frac{dy_m}{dx}\right] - y_m \frac{d}{dx}\left[p(x)\frac{dy_n}{dx}\right] = (\lambda_n - \lambda_m) r(x) y_n y_m. \qquad \textbf{(6.2.4)}$$

Integrating Equation 6.2.4 from a to b yields

$$\int_a^b \left\{ y_n \frac{d}{dx}\left[p(x)\frac{dy_m}{dx}\right] - y_m \frac{d}{dx}\left[p(x)\frac{dy_n}{dx}\right]\right\} dx = (\lambda_n - \lambda_m)\int_a^b r(x) y_n y_m\, dx. \qquad \textbf{(6.2.5)}$$

We can simplify the left side of Equation 6.2.5 by integrating by parts to give

$$\int_a^b \left\{ y_n \frac{d}{dx}\left[p(x)\frac{dy_m}{dx}\right] - y_m \frac{d}{dx}\left[p(x)\frac{dy_n}{dx}\right]\right\} dx$$

$$= \left[p(x) y_m' y_n - p(x) y_n' y_m\right]_a^b - \int_a^b p(x)[y_n' y_m' - y_n' y_m']\, dx. \qquad \textbf{(6.2.6)}$$

The second integral equals zero since the integrand vanishes identically. Because $y_n(x)$ and $y_m(x)$ satisfy the boundary condition at $x = a$,

$$\alpha y_n(a) + \beta y_n'(a) = 0, \qquad \textbf{(6.2.7)}$$

and

$$\alpha y_m(a) + \beta y_m'(a) = 0. \tag{6.2.8}$$

These two equations are simultaneous equations in α and β. Hence, the determinant of the equations must be zero:

$$y_n'(a)y_m(a) - y_m'(a)y_n(a) = 0. \tag{6.2.9}$$

Similarly, at the other end,

$$y_n'(b)y_m(b) - y_m'(b)y_n(b) = 0. \tag{6.2.10}$$

Consequently, the right side of Equation 6.2.6 vanishes and Equation 6.2.5 reduces to Equation 6.2.1. □

• Example 6.2.1

Let us verify the orthogonality condition for the eigenfunctions that we found in Example 6.1.1.

Because $r(x) = 1$, $a = 0$, $b = \pi$, and $y_n(x) = \sin(k_n x)$, we find that

$$\int_a^b r(x) y_n y_m \, dx = \int_0^\pi \sin(k_n x)\sin(k_m x)\, dx \tag{6.2.11}$$

$$= \tfrac{1}{2}\int_0^\pi \{\cos[(k_n - k_m)x] - \cos[(k_n + k_m)x]\}\, dx \tag{6.2.12}$$

$$= \left.\frac{\sin[(k_n - k_m)x]}{2(k_n - k_m)}\right|_0^\pi - \left.\frac{\sin[(k_n + k_m)x]}{2(k_n + k_m)}\right|_0^\pi \tag{6.2.13}$$

$$= \frac{\sin[(k_n - k_m)\pi]}{2(k_n - k_m)} - \frac{\sin[(k_n + k_m)\pi]}{2(k_n + k_m)} \tag{6.2.14}$$

$$= \frac{\sin(k_n\pi)\cos(k_m\pi) - \cos(k_n\pi)\sin(k_m\pi)}{2(k_n - k_m)} - \frac{\sin(k_n\pi)\cos(k_m\pi) + \cos(k_n\pi)\sin(k_m\pi)}{2(k_n + k_m)} \tag{6.2.15}$$

$$= \frac{k_n\cos(k_n\pi)\cos(k_m\pi) - k_m\cos(k_n\pi)\cos(k_m\pi)}{2(k_n - k_m)} - \frac{k_n\cos(k_n\pi)\cos(k_m\pi) + k_m\cos(k_n\pi)\cos(k_m\pi)}{2(k_n + k_m)} \tag{6.2.16}$$

$$= \frac{(k_n - k_m)\cos(k_n\pi)\cos(k_m\pi)}{2(k_n - k_m)} - \frac{(k_n + k_m)\cos(k_n\pi)\cos(k_m\pi)}{2(k_n + k_m)} = 0. \tag{6.2.17}$$

We used the relationships $k_n = \tan(k_n\pi)$, and $k_m = \tan(k_m\pi)$ to simplify Equation 6.2.15. Note, however, that if $n = m$,

$$\int_0^\pi \sin(k_n x)\sin(k_n x)\, dx = \tfrac{1}{2}\int_0^\pi [1 - \cos(2k_n x)]\, dx = \frac{\pi}{2} - \frac{\sin(2k_n\pi)}{4k_n} = \tfrac{1}{2}[\pi - \cos^2(k_n\pi)] > 0, \tag{6.2.18}$$

because $\sin(2A) = 2\sin(A)\cos(A)$, and $k_n = \tan(k_n\pi)$. That is, any eigenfunction *cannot* be orthogonal to itself.

In closing, we note that had we defined the eigenfunction in our example as

$$y_n(x) = \frac{\sin(k_n x)}{\sqrt{[\pi - \cos^2(k_n\pi)]/2}} \tag{6.2.19}$$

rather than $y_n(x) = \sin(k_n x)$, the orthogonality condition would read

$$\int_0^\pi y_n(x)y_m(x)\,dx = \begin{cases} 0, & m \neq n, \\ 1, & m = n. \end{cases} \tag{6.2.20}$$

This process of *normalizing* an eigenfunction so that the orthogonality condition becomes

$$\int_a^b r(x)y_n(x)y_m(x)\,dx = \begin{cases} 0, & m \neq n, \\ 1, & m = n, \end{cases} \tag{6.2.21}$$

generates *orthonormal* eigenfunctions. We will see the convenience of doing this in the next section.

Problems

1. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y(0) = y(L) = 0$ has the eigenfunction solution $y_n(x) = \sin(n\pi x/L)$. By direct integration, verify the orthogonality condition, Equation 6.2.1.

2. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y'(0) = y'(L) = 0$ has the eigenfunction solutions $y_0(x) = 1$ and $y_n(x) = \cos(n\pi x/L)$. By direct integration, verify the orthogonality condition, Equation 6.2.1.

3. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y(0) = y'(L) = 0$ has the eigenfunction solution $y_n(x) = \sin[(2n-1)\pi x/(2L)]$. By direct integration, verify the orthogonality condition, Equation 6.2.1.

4. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y'(0) = y(L) = 0$ has the eigenfunction solution $y_n(x) = \cos[(2n-1)\pi x/(2L)]$. By direct integration, verify the orthogonality condition, Equation 6.2.1.

6.3 EXPANSION IN SERIES OF EIGENFUNCTIONS

In calculus we learned that under certain conditions we could represent a function $f(x)$ by a linear and infinite sum of polynomials $(x - x_0)^n$. In this section we show that an analogous procedure exists for representing a piece-wise continuous function by a linear sum of eigenfunctions. These *eigenfunction expansions* will be used in the next three chapters to solve partial differential equations.

Let the function $f(x)$ be defined in the interval $a < x < b$. We wish to re-express $f(x)$ in terms of the eigenfunctions $y_n(x)$ given by a regular Sturm-Liouville problem. Assuming that the function $f(x)$ can be represented by a uniformly convergent series,[6] we write

$$f(x) = \sum_{n=1}^{\infty} c_n y_n(x). \tag{6.3.1}$$

The orthogonality relation, Equation 6.2.1, gives us the method for computing the coefficients c_n. First we multiply both sides of Equation 6.3.1 by $r(x)y_m(x)$, where m is a fixed integer, and then integrate from a to b. Because this series is uniformly convergent and $y_n(x)$ is continuous, we can integrate the series term by term, or

$$\int_a^b r(x)f(x)y_m(x)\,dx = \sum_{n=1}^{\infty} c_n \int_a^b r(x)y_n(x)y_m(x)\,dx. \tag{6.3.2}$$

The orthogonality relationship states that all of the terms on the right side of Equation 6.3.2 must disappear except the one for which $n = m$. Thus, we are left with

$$\int_a^b r(x)f(x)y_m(x)\,dx = c_m \int_a^b r(x)y_m(x)y_m(x)\,dx \tag{6.3.3}$$

or

$$c_n = \frac{\int_a^b r(x)f(x)y_n(x)\,dx}{\int_a^b r(x)y_n^2(x)\,dx}, \tag{6.3.4}$$

if we replace m by n in Equation 6.3.3.

Usually, both integrals in Equation 6.3.4 are evaluated by direct integration. In the case when the evaluation of the denominator is very difficult, Lockshin[7] has shown that the denominator of Equation 6.3.4 always equals

$$\int_a^b r(x)y^2(x)\,dx = p(x)\left[\frac{\partial y}{\partial x}\frac{\partial y}{\partial \lambda} - y\frac{\partial^2 y}{\partial \lambda \partial x}\right]\Bigg|_a^b, \tag{6.3.5}$$

for a regular Sturm-Liouville problem with eigenfunction solution y, where $p(x)$, $q(x)$, and $r(x)$ are continuously differentiable on the interval $[a, b]$.

The series, Equation 6.3.1, with the coefficients found by Equation 6.3.4, is a *generalized Fourier series* of the function $f(x)$ with respect to the eigenfunction $y_n(x)$. It is called a generalized Fourier series because we generalized the procedure of re-expressing a function $f(x)$ by sines and cosines into one involving solutions to regular Sturm-Liouville problems. Note that if we had used an orthonormal set of eigenfunctions, then the denominator of

---

[6] If $S_n(x) = \sum_{k=1}^{n} u_k(x)$, $S(x) = \lim_{n\to\infty} S_n(x)$, and $0 < |S_n(x) - S(x)| < \epsilon$ for all $n > M > 0$, the series $\sum_{k=1}^{\infty} u_k(x)$ is uniformly convergent if M is dependent on ϵ alone and not x.

[7] Lockshin, J. L, 2001: Explicit closed-form expression for eigenfunction norms. *Appl. Math. Lett.*, **14**, 553–555.

Equation 6.3.4 would equal one and we reduce our work by half. The coefficients c_n are the *Fourier coefficients*.

One of the most remarkable facts about generalized Fourier series is their applicability even when the function has a finite number of bounded discontinuities in the range $[a, b]$. We may formally express this fact by the following theorem:

Theorem: *If both $f(x)$ and $f'(x)$ are piece-wise continuous in $a \le x \le b$, then $f(x)$ can be expanded in a uniformly convergent Fourier series, Equation 6.3.1, whose coefficients c_n are given by Equation 6.3.4. It converges to $[f(x^+) + f(x^-)]/2$ at any point x in the open interval $a < x < b$.*

The proof is beyond the scope of this book but can be found in more advanced treatises.[8] If we are willing to include stronger constraints, we can make even stronger statements about convergence. For example,[9] if we require that $f(x)$ be a continuous function with a piece-wise continuous first derivative, then the eigenfunction expansion, Equation 6.3.1, converges to $f(x)$ uniformly and absolutely in $[a, b]$ if $f(x)$ satisfies the same boundary conditions as does $y_n(x)$. □

In the case when $f(x)$ is discontinuous, we are not merely rewriting $f(x)$ in a new form. We are actually choosing the coefficients c_n so that the eigenfunction expansion fits $f(x)$ in the "least squares" sense that

$$\int_a^b r(x) \left| f(x) - \sum_{n=1}^{\infty} c_n y_n(x) \right|^2 dx = 0. \tag{6.3.6}$$

Consequently we should expect peculiar things, such as spurious oscillations, to occur in the neighborhood of the discontinuity. These are *Gibbs phenomena*,[10] the same phenomena discovered with Fourier series. See Section 5.2.

• Example 6.3.1

To illustrate the concept of an eigenfunction expansion, let us find the expansion for $f(x) = x$ over the interval $0 < x < \pi$ using the solution to the regular Sturm-Liouville problem of

$$y'' + \lambda y = 0, \qquad y(0) = y(\pi) = 0. \tag{6.3.7}$$

This problem arises when we solve the wave or heat equation by separation of variables in the next two chapters.

Because the eigenfunctions are $y_n(x) = \sin(nx)$, $n = 1, 2, 3, \ldots$, $r(x) = 1$, $a = 0$, and $b = \pi$, Equation 6.3.4 yields

$$c_n = \frac{\int_0^\pi x \sin(nx)\, dx}{\int_0^\pi \sin^2(nx)\, dx} = \frac{-x\cos(nx)/n + \sin(nx)/n^2\big|_0^\pi}{x/2 - \sin(2nx)/(4n)\big|_0^\pi} = -\frac{2}{n}\cos(n\pi) = -\frac{2}{n}(-1)^n. \tag{6.3.8}$$

---

[8] For example, Titchmarsh, E. C., 1962: *Eigenfunction Expansions Associated with Second-Order Differential Equations. Part 1*. Oxford University Press, pp. 12–16.

[9] Tolstov, G. P., 1962: *Fourier Series*. Dover Publishers, p. 255.

[10] Apparently first discussed by Weyl, H., 1910: Die Gibbs'sche Erscheinung in der Theorie der Sturm-Liouvilleschen Reihen. *Rend. Circ. Mat. Palermo*, **29**, 321–323.

Equation 6.3.1 then gives

$$f(x) = -2\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin(nx). \qquad \textbf{(6.3.9)}$$

This particular example is in fact an example of a half-range sine expansion.

Finally we must state the values of x for which Equation 6.3.9 is valid. At $x = \pi$ the series converges to zero while $f(\pi) = \pi$. At $x = 0$ both the series and the function converge to zero. Hence this series expansion is valid for $0 \le x < \pi$. □

• Example 6.3.2

For our second example let us find the expansion for $f(x) = x$ over the interval $0 \le x < \pi$ using the solution to the regular Sturm-Liouville problem of

$$y'' + \lambda y = 0, \qquad y(0) = y(\pi) - y'(\pi) = 0. \qquad \textbf{(6.3.10)}$$

We will encounter this problem when we solve the heat equation with radiative boundary conditions by separation of variables.

Because $r(x) = 1$, $a = 0$, $b = \pi$, and the eigenfunctions are $y_n(x) = \sin(k_n x)$, where $k_n = \tan(k_n\pi)$, Equation 6.3.4 yields

$$c_n = \frac{\int_0^\pi x\sin(k_n x)\,dx}{\int_0^\pi \sin^2(k_n x)\,dx} = \frac{\int_0^\pi x\sin(k_n x)\,dx}{\frac{1}{2}\int_0^\pi [1-\cos(2k_n x)]\,dx} = \frac{2\sin(k_n x)/k_n^2 - 2x\cos(k_n x)/k_n\big|_0^\pi}{x - \sin(2k_n x)/(2k_n)\big|_0^\pi} \qquad \textbf{(6.3.11)}$$

$$= \frac{2\sin(k_n\pi)/k_n^2 - 2\pi\cos(k_n\pi)/k_n}{\pi - \sin(2k_n\pi)/(2k_n)} = \frac{2(1-\pi)\cos(k_n\pi)/k_n}{\pi - \cos^2(k_n\pi)}, \qquad \textbf{(6.3.12)}$$

where we used the property that $\sin(k_n\pi) = k_n\cos(k_n\pi)$. Equation 6.3.1 then gives

$$f(x) = 2(1-\pi)\sum_{n=1}^{\infty}\frac{\cos(k_n\pi)}{k_n[\pi - \cos^2(k_n\pi)]}\sin(k_n x). \qquad \textbf{(6.3.13)}$$

To illustrate the use of Equation 6.3.5, we note that

$$y(x) = \sin(\sqrt{\lambda}\,x), \quad \frac{\partial y}{\partial x} = \sqrt{\lambda}\cos(\sqrt{\lambda}\,x), \quad \frac{\partial y}{\partial \lambda} = \frac{x}{2\sqrt{\lambda}}\cos(\sqrt{\lambda}\,x), \qquad \textbf{(6.3.14)}$$

and

$$\frac{\partial^2 y}{\partial\lambda\partial x} = \frac{\partial^2 y}{\partial x\partial\lambda} = \frac{1}{2\sqrt{\lambda}}\cos(\sqrt{\lambda}\,x) - \frac{x}{2}\sin(\sqrt{\lambda}\,x). \qquad \textbf{(6.3.15)}$$

Therefore,

$$\int_0^\pi r(x)y_n^2(x)\,dx = \left\{\frac{x}{2}\cos^2(k_n x) - \sin(k_n x)\left[\frac{1}{2k_n}\cos(k_n x) - \frac{x}{2}\sin(k_n x)\right]\right\}\Bigg|_0^\pi \qquad \textbf{(6.3.16)}$$

$$= \left[\frac{x}{2} - \frac{1}{2k_n}\sin(k_n x)\cos(k_n x)\right]\Bigg|_0^\pi = \frac{\pi}{2} - \frac{\cos^2(k_n\pi)}{2}. \qquad \textbf{(6.3.17)}$$

Note that we set $\lambda = \lambda_n = k_n^2$ after taking the derivatives with respect to λ.

Problems

1. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y(0) = y(L) = 0$ has the eigenfunction solution $y_n(x) = \sin(n\pi x/L)$. Find the eigenfunction expansion for $f(x) = x$ using this eigenfunction.

2. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y'(0) = y'(L) = 0$ has the eigenfunction solutions $y_0(x) = 1$, and $y_n(x) = \cos(n\pi x/L)$. Find the eigenfunction expansion for $f(x) = x$ using these eigenfunctions.

3. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y(0) = y'(L) = 0$ has the eigenfunction solution $y_n(x) = \sin[(2n-1)\pi x/(2L)]$. Find the eigenfunction expansion for $f(x) = x$ using this eigenfunction.

4. The Sturm-Liouville problem $y'' + \lambda y = 0$, $y'(0) = y(L) = 0$ has the eigenfunction solution $y_n(x) = \cos[(2n-1)\pi x/(2L)]$. Find the eigenfunction expansion for $f(x) = x$ using this eigenfunction.

5. Consider the eigenvalue problem

$$y'' + (\lambda - a^2)y = 0, \qquad 0 < x < 1,$$

with the boundary conditions

$$y'(0) + ay(0) = 0 \qquad \text{and} \qquad y'(1) + ay(1) = 0.$$

Step 1: Show that this is a regular Sturm-Liouville problem.
Step 2: Show that the eigenvalues and eigenfunctions are $\lambda_0 = 0$, $y_0(x) = e^{-ax}$ and $\lambda_n = a^2 + n^2\pi^2$, $y_n(x) = a\sin(n\pi x) - n\pi\cos(n\pi x)$.
where $n = 1, 2, 3, \ldots$.
Step 3: Given a function $f(x)$, show that we can expand it as follows:

$$f(x) = C_0 e^{-ax} + \sum_{n=1}^{\infty} C_n \left[a\sin(n\pi x) - n\pi\cos(n\pi x)\right],$$

where

$$\left(1 - e^{-2a}\right) C_0 = 2a \int_0^1 f(x) e^{-ax}\, dx,$$

and

$$(a^2 + n^2\pi^2)C_n = 2\int_0^1 f(x)\left[a\sin(n\pi x) - n\pi\cos(n\pi x)\right] dx.$$

6. Consider the eigenvalue problem

$$y'''' + \lambda y'' = 0, \qquad 0 < x < 1,$$

with the boundary conditions $y(0) = y'(0) = y(1) = y'(1) = 0$. Prove the following points:

Step 1: Show that the eigenfunctions are

$$y_n(x) = 1 - \cos(k_n x) + \frac{1 - \cos(k_n)}{k_n - \sin(k_n)}[\sin(k_n x) - k_n x],$$

where k_n denotes the nth root of

$$2 - 2\cos(k) - k\sin(k) = \sin(k/2)[\sin(k/2) - (k/2)\cos(k/2)] = 0.$$

Step 2: Show that there are two classes of eigenfunctions:

$$\kappa_n = 2n\pi, \qquad y_n(x) = 1 - \cos(2n\pi x),$$

and

$$\tan(\kappa_n/2) = \kappa_n/2, \qquad y_n(x) = 1 - \cos(\kappa_n x) + \frac{2}{\kappa_n}[\sin(\kappa_n x) - \kappa_n x].$$

Step 3: Show that the orthogonality condition for this problem is

$$\int_0^1 y_n'(x) y_m'(x)\, dx = 0, \qquad n \neq m,$$

where $y_n(x)$ and $y_m(x)$ are two distinct eigenfunction solutions of this problem. Hint: Follow the proof in Section 6.2 and integrate repeatedly by parts to eliminate higher derivative terms.

Step 4: Show that we can construct an eigenfunction expansion for an arbitrary function $f(x)$ via

$$f(x) = \sum_{n=1}^{\infty} C_n y_n(x), \qquad 0 < x < 1,$$

provided

$$C_n = \frac{\int_0^1 f'(x) y_n'(x)\, dx}{\int_0^1 [y_n'(x)]^2\, dx}.$$

What are the condition(s) on $f(x)$?

6.4 A SINGULAR STURM-LIOUVILLE PROBLEM: LEGENDRE'S EQUATION

In the previous sections we used solutions to a regular Sturm-Liouville problem in the eigenfunction expansion of the function $f(x)$. The fundamental reason why we could form such an expansion was the orthogonality condition, Equation 6.2.1. This crucial property allowed us to solve for the Fourier coefficient c_n given by Equation 6.3.4.

In the next few chapters, when we solve partial differential equations in cylindrical and spherical coordinates, we will find that $f(x)$ must be expanded in terms of eigenfunctions from singular Sturm-Liouville problems. Is this permissible? How do we compute the Fourier coefficients in this case? The final two sections of this chapter deal with these questions by examining the two most frequently encountered singular Sturm-Liouville problems, those involving Legendre's and Bessel's equations.

Born into an affluent family, Adrien-Marie Legendre's (1752–1833) modest family fortune was sufficient to allow him to devote his life to research in celestial mechanics, number theory, and the theory of elliptic functions. In July 1784 he read before the *Académie des sciences* his *Recherches sur la figure des planètes.* It is in this paper that Legendre polynomials first appeared. (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

We begin by determining the orthogonality condition for singular Sturm-Liouville problems. Returning to the beginning portions of Section 6.2, we combine Equation 6.2.5 and Equation 6.2.6 to obtain

$$(\lambda_n - \lambda_m) \int_a^b r(x) y_n y_m \, dx = [p(b)y'_m(b)y_n(b) - p(b)y'_n(b)y_m(b) - p(a)y'_m(a)y_n(a) + p(a)y'_n(a)y_m(a)]. \quad \textbf{(6.4.1)}$$

From Equation 6.4.1 the right side vanishes and we preserve orthogonality if $y_n(x)$ is finite and $p(x)y'_n(x)$ tends to zero at both endpoints. This is not the only choice but let us see where it leads.

Consider now Legendre's equation:

$$(1 - x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0, \quad \textbf{(6.4.2)}$$

or

$$\frac{d}{dx}\left[(1 - x^2)\frac{dy}{dx}\right] + n(n+1)y = 0, \quad \textbf{(6.4.3)}$$

where we set $a = -1$, $b = 1$, $\lambda = n(n+1)$, $p(x) = 1 - x^2$, $q(x) = 0$, and $r(x) = 1$. This equation arises in the solution of partial differential equations involving spherical geometry. Because $p(-1) = p(1) = 0$, we are faced with a singular Sturm-Liouville problem. Before we can determine if any of its solutions can be used in an eigenfunction expansion, we must find them.

Equation 6.4.2 does not have a simple general solution. [If $n = 0$, then $y(x) = 1$ is a solution.] Consequently we try to solve it with the power series:

$$y(x) = \sum_{k=0}^{\infty} A_k x^k, \tag{6.4.4}$$

$$y'(x) = \sum_{k=0}^{\infty} k A_k x^{k-1}, \tag{6.4.5}$$

and

$$y''(x) = \sum_{k=0}^{\infty} k(k-1) A_k x^{k-2}. \tag{6.4.6}$$

Substituting into Equation 6.4.2,

$$\sum_{k=0}^{\infty} k(k-1) A_k x^{k-2} + \sum_{k=0}^{\infty} \left[n(n+1) - 2k - k(k-1)\right] A_k x^k = 0, \tag{6.4.7}$$

which equals

$$\sum_{m=2}^{\infty} m(m-1) A_m x^{m-2} + \sum_{k=0}^{\infty} \left[n(n+1) - k(k+1)\right] A_k x^k = 0. \tag{6.4.8}$$

If we define $k = m - 2$ in the first summation, then

$$\sum_{k=0}^{\infty} (k+2)(k+1) A_{k+2} x^k + \sum_{k=0}^{\infty} \left[n(n+1) - k(k+1)\right] A_k x^k = 0. \tag{6.4.9}$$

Because Equation 6.4.9 must be true for any x, each power of x must vanish separately. It then follows that

$$(k+2)(k+1) A_{k+2} = [k(k+1) - n(n+1)] A_k, \tag{6.4.10}$$

or

$$A_{k+2} = \frac{[k(k+1) - n(n+1)]}{(k+1)(k+2)} A_k, \tag{6.4.11}$$

where $k = 0, 1, 2, \ldots$. Note that we still have the two arbitrary constants A_0 and A_1 that are necessary for the general solution of Equation 6.4.2.

The first few terms of the solution associated with A_0 are

$$\begin{aligned} u_p(x) = 1 &- \frac{n(n+1)}{2!} x^2 + \frac{n(n-2)(n+1)(n+3)}{4!} x^4 \\ &- \frac{n(n-2)(n-4)(n+1)(n+3)(n+5)}{6!} x^6 + \cdots, \end{aligned} \tag{6.4.12}$$

Table 6.4.1: The First Ten Legendre Polynomials

$$P_0(x) = 1$$
$$P_1(x) = x$$
$$P_2(x) = \tfrac{1}{2}(3x^2 - 1)$$
$$P_3(x) = \tfrac{1}{2}(5x^3 - 3x)$$
$$P_4(x) = \tfrac{1}{8}(35x^4 - 30x^2 + 3)$$
$$P_5(x) = \tfrac{1}{8}(63x^5 - 70x^3 + 15x)$$
$$P_6(x) = \tfrac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5)$$
$$P_7(x) = \tfrac{1}{16}(429x^7 - 693x^5 + 315x^3 - 35x)$$
$$P_8(x) = \tfrac{1}{128}(6435x^8 - 12012x^6 + 6930x^4 - 1260x^2 + 35)$$
$$P_9(x) = \tfrac{1}{128}(12155x^9 - 25740x^7 + 18018x^5 - 4620x^3 + 315x)$$
$$P_{10}(x) = \tfrac{1}{256}(46189x^{10} - 109395x^8 + 90090x^6 - 30030x^4 + 3465x^2 - 63)$$

while the first few terms associated with the A_1 coefficient are

$$v_p(x) = x - \frac{(n-1)(n+2)}{3!}x^3 + \frac{(n-1)(n-3)(n+2)(n+4)}{5!}x^5 - \frac{(n-1)(n-3)(n-5)(n+2)(n+4)(n+6)}{7!}x^7 + \cdots. \quad \textbf{(6.4.13)}$$

If n is an *even* positive integer (including $n = 0$), then the series, Equation 6.4.12, terminates with the term involving x^n: The solution is a polynomial of degree n. Similarly, if n is an *odd* integer, the series, Equation 6.4.13, terminates with the term involving x^n. Otherwise, for n noninteger the expressions are infinite series.

For reasons that will become apparent, we restrict ourselves to positive integers n. Actually, this includes all possible integers because the negative integer $-n-1$ has the same Legendre's equation and solution as the positive integer n. These polynomials are *Legendre polynomials*[11] and we may compute them by the power series:

$$P_n(x) = \sum_{k=0}^{m} (-1)^k \frac{(2n-2k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k}, \quad \textbf{(6.4.14)}$$

where $m = n/2$, or $m = (n-1)/2$, depending upon which is an integer. We chose to use Equation 6.4.14 over Equation 6.4.12 or Equation 6.4.13 because Equation 6.4.14 has the advantage that $P_n(1) = 1$. Table 6.4.1 gives the first ten Legendre polynomials.

The other solution, the infinite series, is the Legendre function of the second kind, $Q_n(x)$. Figure 6.4.1 illustrates the first four Legendre polynomials $P_n(x)$ while Figure 6.4.2

[11] Legendre, A. M., 1785: Sur l'attraction des sphéroïdes homogénes. *Mém. math. phys. présentés à l'Acad. sci. pars divers savants*, **10**, 411–434. The best reference on Legendre polynomials is Hobson, E. W., 1965: *The Theory of Spherical and Ellipsoidal Harmonics.* Chelsea Publishing Co., 500 pp.

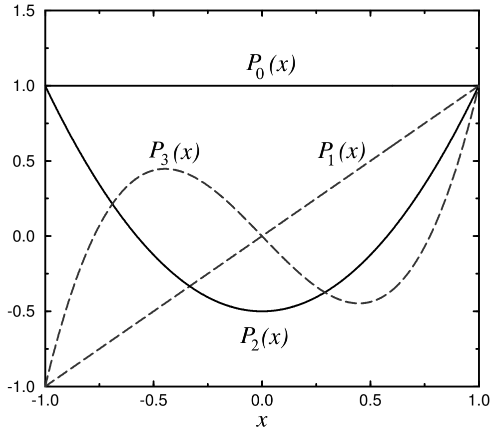


Figure 6.4.1: The first four Legendre functions of the first kind.

gives the first four Legendre functions of the second kind $Q_n(x)$. From this figure we see that $Q_n(x)$ becomes infinite at the points $x = \pm 1$. As shown earlier, this is important because we are only interested in solutions to Legendre's equation that are finite over the interval $[-1, 1]$. On the other hand, in problems where we exclude the points $x = \pm 1$, Legendre functions of the second kind will appear in the general solution.[12]

In the case that n is not an integer, we can construct a solution[13] that remains finite at $x = 1$ but not at $x = -1$. Furthermore, we can construct a solution that is finite at $x = -1$ but not at $x = 1$. Because our solutions must be finite at both endpoints so that we can use them in an eigenfunction expansion, we must reject these solutions from further consideration and are left only with Legendre polynomials. From now on, we will only consider the properties and uses of these polynomials.

Although we have the series, Equation 6.4.14, to compute $P_n(x)$, there are several alternative methods. We obtain the first method, known as *Rodrigues's formula*,[14] by writing Equation 6.4.14 in the form

$$P_n(x) = \frac{1}{2^n n!} \sum_{k=0}^{n} (-1)^k \frac{n!}{k!(n-k)!} \frac{(2n-2k)!}{(n-2k)!} x^{n-2k} \qquad \textbf{(6.4.15)}$$

$$= \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[\sum_{k=0}^{n} (-1)^k \frac{n!}{k!(n-k)!} x^{2n-2k} \right]. \qquad \textbf{(6.4.16)}$$

The last summation is the binomial expansion of $(x^2 - 1)^n$ so that

$$\boxed{P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.} \qquad \textbf{(6.4.17)}$$

[12] See Smythe, W. R., 1950: *Static and Dynamic Electricity*. McGraw-Hill, Section 5.215, for an example.

[13] See Carrier, G. F., M. Krook, and C. E. Pearson, 1966: *Functions of the Complex Variable: Theory and Technique*. McGraw-Hill, pp. 212–213.

[14] Rodrigues, O., 1816: Mémoire sur l'attraction des sphéroïdes. *Correspond. l'École Polytech.*, **3**, 361–385.

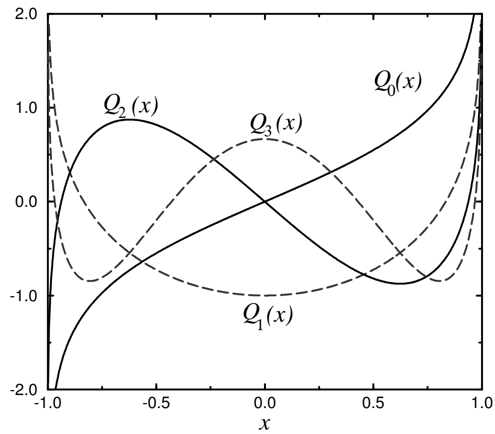


Figure 6.4.2: The first four Legendre functions of the second kind.

Another method for computing $P_n(x)$ involves the use of recurrence formulas. The first step in finding these formulas is to establish the fact that

$$(1 + h^2 - 2xh)^{-1/2} = P_0(x) + hP_1(x) + h^2P_2(x) + \cdots. \qquad \textbf{(6.4.18)}$$

The function $(1 + h^2 - 2xh)^{-1/2}$ is the *generating function* for $P_n(x)$. We obtain the expansion via the formal binomial expansion

$$(1 + h^2 - 2xh)^{-1/2} = 1 + \tfrac{1}{2}(2xh - h^2) + \tfrac{1}{2}\tfrac{3}{2}\tfrac{1}{2!}(2xh - h^2)^2 + \cdots. \qquad \textbf{(6.4.19)}$$

Upon expanding the terms contained in $2x - h^2$ and grouping like powers of h,

$$(1 + h^2 - 2xh)^{-1/2} = 1 + xh + (\tfrac{3}{2}x^2 - \tfrac{1}{2})h^2 + \cdots. \qquad \textbf{(6.4.20)}$$

A direct comparison between the coefficients of each power of h and the Legendre polynomial $P_n(x)$ completes the demonstration. Note that these results hold only if $|x|$ and $|h| < 1$.

Next we define $W(x, h) = (1 + h^2 - 2xh)^{-1/2}$. A quick check shows that $W(x, h)$ satisfies the first-order partial differential equation

$$(1 - 2xh + h^2)\frac{\partial W}{\partial h} + (h - x)W = 0. \qquad \textbf{(6.4.21)}$$

The substitution of Equation 6.4.18 into Equation 6.4.21 yields

$$(1 - 2xh + h^2)\sum_{n=0}^{\infty} nP_n(x)h^{n-1} + (h - x)\sum_{n=0}^{\infty} P_n(x)h^n = 0. \qquad \textbf{(6.4.22)}$$

Setting the coefficients of h^n equal to zero, we find that

$$(n + 1)P_{n+1}(x) - 2nxP_n(x) + (n - 1)P_{n-1}(x) + P_{n-1}(x) - xP_n(x) = 0, \qquad \textbf{(6.4.23)}$$

or

$$\boxed{(n + 1)P_{n+1}(x) - (2n + 1)xP_n(x) + nP_{n-1}(x) = 0} \qquad \textbf{(6.4.24)}$$

with $n = 1, 2, 3, \ldots$.

Similarly, the first-order partial differential equation

$$(1-2xh+h^2)\frac{\partial W}{\partial x} - hW = 0 \tag{6.4.25}$$

leads to

$$(1-2xh+h^2)\sum_{n=0}^{\infty} P_n'(x)h^n - \sum_{n=0}^{\infty} P_n(x)h^{n+1} = 0, \tag{6.4.26}$$

which implies

$$P_{n+1}'(x) - 2xP_n'(x) + P_{n-1}'(x) - P_n(x) = 0. \tag{6.4.27}$$

Differentiating Equation 6.4.24, we first eliminate $P_{n-1}'(x)$ and then $P_{n+1}'(x)$ from the resulting equations and Equation 6.4.27. This gives two further recurrence relationships:

$$P_{n+1}'(x) - xP_n'(x) - (n+1)P_n(x) = 0, \quad n = 0,1,2,\ldots, \tag{6.4.28}$$

and

$$xP_n'(x) - P_{n-1}'(x) - nP_n(x) = 0, \quad n = 1,2,3,\ldots. \tag{6.4.29}$$

Adding Equation 6.4.28 and Equation 6.4.29, we obtain the more symmetric formula

$$P_{n+1}'(x) - P_{n-1}'(x) = (2n+1)P_n(x),\ n = 1,2,3,\ldots. \tag{6.4.30}$$

Given any two of the polynomials $P_{n+1}(x)$, $P_n(x)$, and $P_{n-1}(x)$, Equation 6.4.24 or Equation 6.4.30 yields the third.

Having determined several methods for finding the Legendre polynomial $P_n(x)$, we now turn to the actual orthogonality condition.[15] Consider the integral

$$J = \int_{-1}^{1} \frac{dx}{\sqrt{1+h^2-2xh}\,\sqrt{1+t^2-2xt}}, \qquad |h|,|t|<1 \tag{6.4.31}$$

$$= \int_{-1}^{1} [P_0(x) + hP_1(x) + \cdots + h^n P_n(x) + \cdots] \times [P_0(x) + tP_1(x) + \cdots + t^n P_n(x) + \cdots]\,dx \tag{6.4.32}$$

$$= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} h^n t^m \int_{-1}^{1} P_n(x)P_m(x)\,dx. \tag{6.4.33}$$

On the other hand, if $a = (1+h^2)/2h$, and $b = (1+t^2)/2t$, the integral J is

$$J = \int_{-1}^{1} \frac{dx}{\sqrt{1+h^2-2xh}\,\sqrt{1+t^2-2xt}} \tag{6.4.34}$$

$$= \frac{1}{2\sqrt{ht}} \int_{-1}^{1} \frac{dx}{\sqrt{a-x}\,\sqrt{b-x}} = \frac{1}{\sqrt{ht}} \int_{-1}^{1} \frac{\frac{1}{2}\left(\frac{1}{\sqrt{a-x}} + \frac{1}{\sqrt{b-x}}\right)}{\sqrt{a-x}+\sqrt{b-x}}\,dx \tag{6.4.35}$$

$$= -\frac{1}{\sqrt{ht}} \ln\left(\sqrt{a-x}+\sqrt{b-x}\right)\Big|_{-1}^{1} = \frac{1}{\sqrt{ht}} \ln\left(\frac{\sqrt{a+1}+\sqrt{b+1}}{\sqrt{a-1}+\sqrt{b-1}}\right). \tag{6.4.36}$$

---

[15] See Symons, B., 1982: Legendre polynomials and their orthogonality. *Math. Gaz.*, **66**, 152–154.

Some Useful Relationships Involving Legendre Polynomials

Rodrigues's formula

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$$

Recurrence formulas

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0, \qquad n = 1, 2, 3, \ldots$$

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)P_n(x), \qquad n = 1, 2, 3, \ldots$$

Orthogonality condition

$$\int_{-1}^{1} P_n(x)P_m(x)\,dx = \begin{cases} 0, & m \neq n, \\ \dfrac{2}{2n+1}, & m = n. \end{cases}$$

But $a+1 = (1+h^2+2h)/2h = (1+h)^2/2h$, and $a-1 = (1-h)^2/2h$. After a little algebra,

$$J = \frac{1}{\sqrt{ht}}\ln\left(\frac{1+\sqrt{ht}}{1-\sqrt{ht}}\right) = \frac{2}{\sqrt{ht}}\left[\sqrt{ht} + \frac{1}{3}\sqrt{(ht)^3} + \frac{1}{5}\sqrt{(ht)^5} + \cdots\right] \quad \textbf{(6.4.37)}$$

$$= 2\left(1 + \frac{ht}{3} + \frac{h^2t^2}{5} + \cdots + \frac{h^nt^n}{2n+1} + \cdots\right). \quad \textbf{(6.4.38)}$$

As we noted earlier, the coefficient of h^nt^m in this series is $\int_{-1}^{1} P_n(x)P_m(x)\,dx$. If we match the powers of h^nt^m, the orthogonality condition is

$$\boxed{\int_{-1}^{1} P_n(x)P_m(x)\,dx = \begin{cases} 0, & m \neq n, \\ \frac{2}{2n+1}, & m = n. \end{cases}} \quad \textbf{(6.4.39)}$$

With the orthogonality condition, Equation 6.4.39, we are ready to show that we can represent a function $f(x)$, which is piece-wise differentiable in the interval $(-1, 1)$, by the series:

$$\boxed{f(x) = \sum_{m=0}^{\infty} A_m P_m(x), \qquad -1 \leq x \leq 1.} \quad \textbf{(6.4.40)}$$

To find A_m we multiply both sides of Equation 6.4.40 by $P_n(x)$ and integrate from -1 to 1:

$$\int_{-1}^{1} f(x)P_n(x)\,dx = \sum_{m=0}^{\infty} A_m \int_{-1}^{1} P_n(x)P_m(x)\,dx. \quad \textbf{(6.4.41)}$$

All of the terms on the right side vanish except for $n = m$ because of the orthogonality condition, Equation 6.4.39. Consequently, the coefficient A_n is

$$A_n \int_{-1}^{1} P_n^2(x)\, dx = \int_{-1}^{1} f(x) P_n(x)\, dx, \tag{6.4.42}$$

or

$$A_n = \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x)\, dx. \tag{6.4.43}$$

In the special case when $f(x)$ and its first n derivatives are continuous throughout the interval $(-1, 1)$, we may use Rodrigues' formula to evaluate

$$\int_{-1}^{1} f(x) P_n(x)\, dx = \frac{1}{2^n n!} \int_{-1}^{1} f(x) \frac{d^n(x^2-1)^n}{dx^n}\, dx = \frac{(-1)^n}{2^n n!} \int_{-1}^{1} (x^2-1)^n f^{(n)}(x)\, dx \tag{6.4.44}$$

by integrating by parts n times. Consequently,

$$A_n = \frac{2n+1}{2^{n+1} n!} \int_{-1}^{1} (1-x^2)^n f^{(n)}(x)\, dx. \tag{6.4.45}$$

A particularly useful result follows from Equation 6.4.45 if $f(x)$ is a polynomial of degree k. Because all derivatives of $f(x)$ of order n vanish identically when $n > k$, $A_n = 0$ if $n > k$. Consequently, any polynomial of degree k can be expressed as a linear combination of the first $k+1$ Legendre polynomials $[P_0(x), \ldots, P_k(x)]$. Another way of viewing this result is to recognize that any polynomial of degree k is an expansion in powers of x. When we expand in Legendre polynomials we are merely regrouping these powers of x into new groups that can be identified as $P_0(x), P_1(x), P_2(x), \ldots, P_k(x)$.

• Example 6.4.1

Let us use Rodrigues' formula to compute $P_2(x)$. From Equation 6.4.17 with $n = 2$,

$$P_2(x) = \frac{1}{2^2 2!} \frac{d^2}{dx^2}[(x^2-1)^2] = \frac{1}{8} \frac{d^2}{dx^2}(x^4 - 2x^2 - 1) = \frac{1}{2}(3x^2 - 1). \tag{6.4.46}$$

□

• Example 6.4.2

Let us compute $P_3(x)$ from a recurrence relation. From Equation 6.4.24 with $n = 2$,

$$3P_3(x) - 5xP_2(x) + 2P_1(x) = 0. \tag{6.4.47}$$

But $P_2(x) = (3x^2 - 1)/2$, and $P_1(x) = x$, so that

$$3P_3(x) = 5xP_2(x) - 2P_1(x) = 5x[(3x^2-1)/2] - 2x = \tfrac{15}{2}x^3 - \tfrac{9}{2}x, \tag{6.4.48}$$

or

$$P_3(x) = (5x^3 - 3x)/2. \tag{6.4.49}$$

□

• Example 6.4.3

We want to show that

$$\int_{-1}^{1} P_n(x)\,dx = 0, \qquad n > 0. \tag{6.4.50}$$

From Equation 6.4.30,

$$(2n+1)\int_{-1}^{1} P_n(x)\,dx = \int_{-1}^{1} [P'_{n+1}(x) - P'_{n-1}(x)]\,dx \tag{6.4.51}$$

$$= P_{n+1}(x) - P_{n-1}(x)\Big|_{-1}^{1} \tag{6.4.52}$$

$$= P_{n+1}(1) - P_{n-1}(1) - P_{n+1}(-1) + P_{n-1}(-1) = 0, \tag{6.4.53}$$

because $P_n(1) = 1$ and $P_n(-1) = (-1)^n$. □

• Example 6.4.4

Let us express $f(x) = x^2$ in terms of Legendre polynomials. The results from Equation 6.4.45 mean that we need only worry about $P_0(x)$, $P_1(x)$, and $P_2(x)$:

$$x^2 = A_0P_0(x) + A_1P_1(x) + A_2P_2(x). \tag{6.4.54}$$

Substituting for the Legendre polynomials,

$$x^2 = A_0 + A_1x + \tfrac{1}{2}A_2(3x^2 - 1), \tag{6.4.55}$$

and

$$A_0 = \tfrac{1}{3}, \quad A_1 = 0, \quad \text{and} \quad A_2 = \tfrac{2}{3}. \tag{6.4.56}$$

□

• Example 6.4.5

Let us find the expansion in Legendre polynomials of the function:

$$f(x) = \begin{cases} 0, & -1 < x < 0, \\ 1, & 0 < x < 1. \end{cases} \tag{6.4.57}$$

We could have done this expansion as a Fourier series but in the solution of partial differential equations on a sphere we must make the expansion in Legendre polynomials.

In this problem, we find that

$$A_n = \frac{2n+1}{2}\int_0^1 P_n(x)\,dx. \tag{6.4.58}$$

Therefore,

$$A_0 = \tfrac{1}{2}\int_0^1 1\,dx = \tfrac{1}{2}, \qquad A_1 = \tfrac{3}{2}\int_0^1 x\,dx = \tfrac{3}{4}, \tag{6.4.59}$$

$$A_2 = \tfrac{5}{2}\int_0^1 \tfrac{1}{2}(3x^2-1)\,dx = 0, \quad \text{and} \quad A_3 = \tfrac{7}{2}\int_0^1 \tfrac{1}{2}(5x^3-3x)\,dx = -\tfrac{7}{16}, \tag{6.4.60}$$

so that

$$f(x) = \tfrac{1}{2}P_0(x) + \tfrac{3}{4}P_1(x) - \tfrac{7}{16}P_3(x) + \tfrac{11}{32}P_5(x) + \cdots. \tag{6.4.61}$$

Figure 6.4.3 illustrates the expansion, Equation 6.4.61, where we used only the first four terms. It was created using the MATLAB script

```
clear;
x = [-1:0.01:1]; % create x points in plot
f = zeros(size(x)); % initialize function f(x)
for k = 1:length(x) % construct function f(x)
  if x(k) < 0; f(k) = 0; else f(k) = 1; end;
end
% initialize Fourier-Legendre series with zeros
flegendre = zeros(size(x));
% read in Fourier coefficients
a(1) = 1/2; a(2) = 3/4; a(3) = 0;
a(4) = -7/16; a(5) = 0; a(6) = 11/32;
clf % clear any figures
for n = 1:6
% compute Legendre polynomial
  N = n-1; P = legendre(N,x);
% compute Fourier-Legendre series
  flegendre = flegendre + a(n) * P(1,:);
% create plot of truncated Fourier-Legendre series
%     with n terms
  if n==1 subplot(2,2,1), plot(x,flegendre,x,f,'--');
    legend('one term','f(x)'); legend boxoff; end
  if n==2 subplot(2,2,2), plot(x,flegendre,x,f,'--');
    legend('two terms','f(x)'); legend boxoff; end
  if n==4 subplot(2,2,3), plot(x,flegendre,x,f,'--');
    legend('four terms','f(x)'); legend boxoff;
    xlabel('x','Fontsize',20); end
  if n==6 subplot(2,2,4), plot(x,flegendre,x,f,'--');
    legend('six terms','f(x)'); legend boxoff;
    xlabel('x','Fontsize',20); end
  axis([-1 1 -0.5 1.5])
end
```

As we add each additional term in the orthogonal expansion, the expansion fits $f(x)$ better in the "least squares" sense of Equation 6.3.5. The spurious oscillations arise from trying to represent a discontinuous function by four continuous, oscillatory functions. Even if

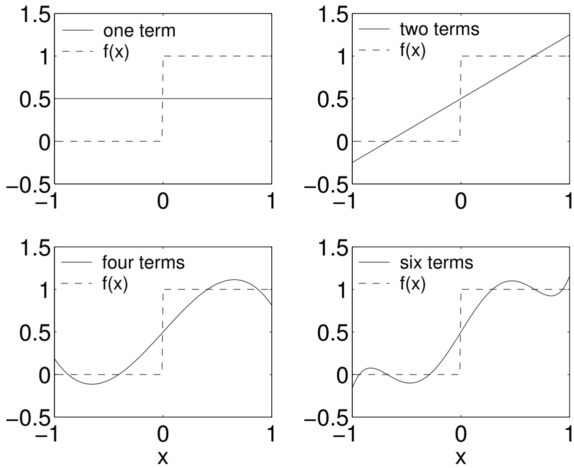


Figure 6.4.3: Representation of the function $f(x) = 1$ for $0 < x < 1$ and 0 for $-1 < x < 0$ by various partial summations of its Legendre polynomial expansion. The dashed lines denote the exact function.

we add additional terms, the spurious oscillations persist, although located nearer to the discontinuity. This is another example of *Gibbs phenomena*.[16] See Section 5.2. □

• Example 6.4.6: Iterative solution of the radiative transfer equation

One of the fundamental equations of astrophysics is the integro-differential equation that describes radiative transfer (the propagation of energy by radiative, rather than conductive or convective, processes) in a gas.

Consider a gas that varies in only one spatial direction and that we divide into infinitesimally thin slabs. As radiation enters a slab, it is absorbed and scattered. If we assume that all of the radiation undergoes isotropic scattering, the radiative transfer equation is

$$\mu \frac{dI}{d\tau} = I - \tfrac{1}{2} \int_{-1}^{1} I \, d\mu, \tag{6.4.62}$$

where I is the intensity of the radiation, τ is the optical depth (a measure of the absorbing power of the gas and related to the distance that you travel within the gas), $\mu = \cos(\theta)$, and θ is the angle at which radiation enters the slab. In this example, we show how the Fourier-Legendre expansion[17]

$$I(\tau, \mu) = \sum_{n=0}^{\infty} I_n(\tau) P_n(\mu) \tag{6.4.63}$$

may be used to solve Equation 6.4.62. Here $I_n(\tau)$ is the Fourier coefficient in the Fourier-Legendre expansion involving the Legendre polynomial $P_n(\mu)$.

We begin by substituting Equation 6.4.63 into Equation 6.4.62,

$$\sum_{n=0}^{\infty} \frac{[(n+1)P_{n+1}(\mu) + nP_{n-1}(\mu)]}{2n+1} \frac{dI_n}{d\tau} = \sum_{n=0}^{\infty} I_n P_n(\mu) - I_0, \tag{6.4.64}$$

---

[16] Weyl, H., 1910: Die Gibbs'sche Erscheinung in der Theorie der Kugelfunktionen. *Rend. Circ. Mat. Palermo*, **29**, 308–321.

[17] See Chandrasekhar, S., 1944: On the radiative equilibrium of a stellar atmosphere. *Astrophys. J.*, **99**, 180–190.

where we used Equation 6.4.24 to eliminate $\mu P_n(\mu)$. Note that only the $I_0(\tau)$ term remains after integrating because of the orthogonality condition:

$$\int_{-1}^{1} 1 \cdot P_n(\mu)\, d\mu = \int_{-1}^{1} P_0(\mu) P_n(\mu)\, d\mu = 0, \tag{6.4.65}$$

if $n > 0$. Equating the coefficients of the various Legendre polynomials,

$$\frac{n}{2n-1} \frac{dI_{n-1}}{d\tau} + \frac{n+1}{2n+3} \frac{dI_{n+1}}{d\tau} = I_n, \tag{6.4.66}$$

for $n = 1, 2, \ldots$ and

$$\frac{dI_1}{d\tau} = 0. \tag{6.4.67}$$

Thus, the solution for I_1 is I_1 = constant = $3F/4$, where F is the net integrated flux and an observable quantity.

For $n = 1$,

$$\frac{dI_0}{d\tau} + \frac{2}{5} \frac{dI_2}{d\tau} = I_1 = \frac{3F}{4}. \tag{6.4.68}$$

Therefore,

$$I_0 + \tfrac{2}{5} I_2 = \tfrac{3}{4} F\tau + A. \tag{6.4.69}$$

The next differential equation arises from $n = 2$ and equals

$$\frac{2}{3} \frac{dI_1}{d\tau} + \frac{3}{7} \frac{dI_3}{d\tau} = I_2. \tag{6.4.70}$$

Because I_1 is a constant and we only retain I_0, I_1, and I_2 in the simplest approximation, we neglect $dI_3/d\tau$ and $I_2 = 0$. Thus, the simplest approximate solution is

$$I_0 = \tfrac{3}{4} F\tau + A, \quad I_1 = \tfrac{3}{4} F, \quad \text{and} \quad I_2 = 0. \tag{6.4.71}$$

To complete our approximate solution, we must evaluate A. If we are dealing with a stellar atmosphere where we assume no external radiation incident on the star, $I(0, \mu) = 0$ for $-1 \le \mu < 0$. Therefore,

$$\int_{-1}^{1} I(\tau, \mu) P_n(\mu)\, d\mu = \sum_{m=0}^{\infty} I_m(\tau) \int_{-1}^{1} P_m(\mu) P_n(\mu)\, d\mu = \frac{2}{2n+1} I_n(\tau). \tag{6.4.72}$$

Taking the limit $\tau \to 0$ and using the boundary condition,

$$\frac{2}{2n+1} I_n(0) = \int_0^1 I(0, \mu) P_n(\mu)\, d\mu = \sum_{m=0}^{\infty} I_m(0) \int_0^1 P_n(\mu) P_m(\mu)\, d\mu. \tag{6.4.73}$$

Thus, we must satisfy, in principle, an infinite set of equations. For example, for $n = 0$, 1, and 2,

$$2I_0(0) = I_0(0) + \tfrac{1}{2} I_1(0) - \tfrac{1}{8} I_3(0) + \tfrac{1}{16} I_5(0) + \cdots, \tag{6.4.74}$$

$$\tfrac{2}{3} I_1(0) = \tfrac{1}{2} I_0(0) + \tfrac{1}{3} I_1(0) + \tfrac{1}{8} I_2(0) - \tfrac{1}{48} I_4(0) + \cdots, \tag{6.4.75}$$

and

$$\tfrac{2}{5} I_2(0) = \tfrac{1}{8} I_1(0) + \tfrac{1}{5} I_2(0) + \tfrac{1}{8} I_3(0) - \tfrac{5}{128} I_5(0) + \cdots. \tag{6.4.76}$$

Using $I_1(0) = 3F/4$,

$$\tfrac{1}{2}I_0(0) + \tfrac{1}{16}I_3(0) - \tfrac{1}{32}I_5(0) + \cdots = \tfrac{3}{16}F, \qquad (\mathbf{6.4.77})$$

$$\tfrac{1}{2}I_0(0) + \tfrac{1}{8}I_2(0) - \tfrac{1}{48}I_4(0) + \cdots = \tfrac{1}{4}F, \qquad (\mathbf{6.4.78})$$

and

$$\tfrac{2}{5}I_2(0) - \tfrac{1}{4}I_3(0) + \tfrac{5}{64}I_5(0) + \cdots = \tfrac{3}{16}F. \qquad (\mathbf{6.4.79})$$

Of the two possible equations, Equation 6.4.77 or Equation 6.4.78, Chandrasekhar chose Equation 6.4.78 from physical considerations. Thus, to first approximation, the solution is

$$I(\mu,\tau) = \tfrac{3}{4}F\left(\tau + \tfrac{2}{3}\right) + \tfrac{3}{4}F\mu + \cdots. \qquad (\mathbf{6.4.80})$$

Better approximations can be obtained by including more terms; the interested reader is referred to the original article. In the early 1950s, Wang and Guth[18] improved the procedure for finding the successive approximations and formulating the approximate boundary conditions.

Problems

Find the first three nonvanishing coefficients in the Legendre polynomial expansion for the following functions:

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

2. $f(x) = \begin{cases} 1/(2\epsilon), & |x| < \epsilon, \\ 0, & \epsilon < |x| < 1, \\ x, & 0 < x < 1. \end{cases}$

3. $f(x) = |x|, \qquad |x| < 1.$

4. $f(x) = x^3, \qquad |x| < 1.$

5. $f(x) = \begin{cases} -1, & -1 < x < 0, \\ 1, & 0 < x < 1. \end{cases}$

6. $f(x) = \begin{cases} -1, & -1 < x < 0, \\ x, & 0 < x < 1. \end{cases}$

Then use MATLAB to illustrate various partial sums of the Fourier-Legendre series.

7. Use Rodrigues' formula to show that $P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$.

8. Given $P_5(x) = \frac{63}{8}x^5 - \frac{70}{8}x^3 + \frac{15}{8}x$ and $P_4(x)$ from Problem 7, use the recurrence formula for $P_{n+1}(x)$ to find $P_6(x)$.

9. Show that (a) $P_n(1) = 1$, (b) $P_n(-1) = (-1)^n$, (c) $P_{2n+1}(0) = 0$, and (d) $P_{2n}(0) = (-1)^n(2n)!/(2^{2n}n!n!)$.

10. Prove that

$$\int_x^1 P_n(t)\,dt = \frac{1}{2n+1}[P_{n-1}(x) - P_{n+1}(x)], \qquad n > 0.$$

---

[18] Wang, M. C., and E. Guth, 1951: On the theory of multiple scattering, particularly of charged particles. *Phys. Rev., Ser. 2*, **84**, 1092–1111.

11. Given[19]

$$P_n[\cos(\theta)] = \frac{2}{\pi}\int_0^\theta \frac{\cos[(n+\frac{1}{2})x]}{\sqrt{2[\cos(x)-\cos(\theta)]}}\,dx = \frac{2}{\pi}\int_\theta^\pi \frac{\sin[(n+\frac{1}{2})x]}{\sqrt{2[\cos(\theta)-\cos(x)]}}\,dx,$$

show that the following generalized Fourier series holds:

$$\frac{H(\theta-t)}{\sqrt{2\cos(t)-2\cos(\theta)}} = \sum_{n=0}^{\infty} P_n[\cos(\theta)]\cos\left[\left(n+\tfrac{1}{2}\right)t\right], \quad 0 \le t < \theta \le \pi,$$

if we use the eigenfunction $y_n(x) = \cos\left[\left(n+\frac{1}{2}\right)x\right]$, $0 < x < \pi$, $r(x) = 1$ and $H(\cdot)$ is Heaviside's step function, and

$$\frac{H(t-\theta)}{\sqrt{2\cos(\theta)-2\cos(t)}} = \sum_{n=0}^{\infty} P_n[\cos(\theta)]\sin\left[\left(n+\tfrac{1}{2}\right)t\right], \quad 0 \le \theta < t \le \pi,$$

if we use the eigenfunction $y_n(x) = \sin\left[\left(n+\frac{1}{2}\right)x\right]$, $0 < x < \pi$, $r(x) = 1$ and $H(\cdot)$ is Heaviside's step function.

12. The series given in Problem 11 are also expansions in Legendre polynomials. In that light, show that

$$\int_0^t \frac{P_n[\cos(\theta)]\,\sin(\theta)}{\sqrt{2\cos(\theta)-2\cos(t)}}\,d\theta = \frac{\sin\left[\left(n+\frac{1}{2}\right)t\right]}{n+\frac{1}{2}},$$

and

$$\int_t^\pi \frac{P_n[\cos(\theta)]\,\sin(\theta)}{\sqrt{2\cos(t)-2\cos(\theta)}}\,d\theta = \frac{\cos\left[\left(n+\frac{1}{2}\right)t\right]}{n+\frac{1}{2}},$$

where $0 < t < \pi$.

13. (a) Use the generating function, Equation 6.4.18, to show that

$$\frac{1}{\sqrt{1-2tx+t^2}} = \sum_{n=0}^{\infty} t^{-n-1}P_n(x), \qquad |x| < 1, \; 1 < |t|.$$

(b) Use the results from part (a) to show that

$$\frac{1}{\sqrt{\cosh(\mu)-x}} = \sqrt{2}\sum_{n=0}^{\infty} e^{-(n+\frac{1}{2})|\mu|}P_n(x), \qquad |x| < 1.$$

Hint:

$$\frac{1}{\sqrt{\cosh(\mu)-x}} = \frac{\sqrt{2}}{\sqrt{e^{|\mu|}-2x+e^{-|\mu|}}}.$$

[19] Hobson, E. W., 1965: *The Theory of Spherical and Ellipsoidal Harmonics*. Chelsea Publishing Co., pp. 26–27.

14. The generating function, Equation 6.4.18, actually holds[20] for $|h| \le 1$ if $|x| < 1$. Using this relationship, show that

$$\sum_{n=0}^{\infty} P_n(x) = \frac{1}{\sqrt{2(1-x)}}, \qquad |x| < 1,$$

and

$$\sum_{n=0}^{\infty} \frac{P_n(x)}{n+1} = \ln\left[\frac{1+\sqrt{(1-x)/2}}{\sqrt{(1-x)/2}}\right], \qquad |x| < 1.$$

Use these relationships to show that

$$\sum_{n=1}^{\infty} \frac{2n+1}{n+1} P_n(x) = 2\sum_{n=1}^{\infty} P_n(x) - \sum_{n=1}^{\infty} \frac{P_n(x)}{n+1} = \frac{1}{\sqrt{(1-x)/2}} - \ln\left[\frac{1+\sqrt{(1-x)/2}}{\sqrt{(1-x)/2}}\right] - 1,$$

if $|x| < 1$.

6.5 ANOTHER SINGULAR STURM-LIOUVILLE PROBLEM: BESSEL'S EQUATION

In the previous section we discussed the solutions to Legendre's equation, especially with regard to their use in orthogonal expansions. In this section we consider another classic equation, Bessel's equation[21]

$$x^2y'' + xy' + (\mu^2x^2 - n^2)y = 0, \tag{6.5.1}$$

or

$$\frac{d}{dx}\left(x\frac{dy}{dx}\right) + \left(\mu^2 x - \frac{n^2}{x}\right)y = 0. \tag{6.5.2}$$

Once again, our ultimate goal is the use of its solutions in orthogonal expansions. These orthogonal expansions, in turn, are used in the solution of partial differential equations in cylindrical coordinates.

A quick check of Bessel's equation shows that it conforms to the canonical form of the Sturm-Liouville problem: $p(x) = x$, $q(x) = -n^2/x$, $r(x) = x$, and $\lambda = \mu^2$. Restricting our attention to the interval $[0, L]$, the Sturm-Liouville problem involving Equation 6.5.2 is singular because $p(0) = 0$. From Equation 6.4.1 in the previous section, the eigenfunctions to a singular Sturm-Liouville problem will still be orthogonal over the interval $[0, L]$ if (1) $y(x)$ is finite and $xy'(x)$ is zero at $x = 0$, and (2) $y(x)$ satisfies the homogeneous boundary condition, Equation 6.1.2, at $x = L$. Consequently, we only seek solutions that satisfy these conditions.

We cannot write down the solution to Bessel's equation in a simple closed form; as in the case with Legendre's equation, we must find the solution by power series. Because we intend to make the expansion about $x = 0$ and this point is a regular singular point, we must

[20] Ibid., p. 28.

[21] Bessel, F. W., 1824: Untersuchung des Teils der planetarischen Störungen, welcher aus der Bewegung der Sonne entsteht. *Abh. d. K. Akad. Wiss. Berlin*, 1–52. See Dutka, J., 1995: On the early history of Bessel functions. *Arch. Hist. Exact Sci.*, **49**, 105–134. The classic reference on Bessel functions is Watson, G. N., 1966: *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, 804 pp.

It was Friedrich Wilhelm Bessel's (1784–1846) apprenticeship to the famous mercantile firm of Kulenkamp that ignited his interest in mathematics and astronomy. As the founder of the German school of practical astronomy, Bessel discovered his functions while studying the problem of planetary motion. Bessel functions arose as coefficients in one of the series that described the gravitational interaction between the sun and two other planets in elliptic orbit. (Portrait courtesy of Photo AKG, London, with permission.)

use the method of Frobenius, where n is an integer.[22] Moreover, because the quantity n^2 appears in Equation 6.5.2, we may take n to be nonnegative without any loss of generality.

To simplify matters, we first find the solution when $\mu = 1$; the solution for $\mu \neq 1$ follows by substituting μx for x. Consequently, we seek solutions of the form

$$y(x) = \sum_{k=0}^{\infty} B_k x^{2k+s}, \tag{6.5.3}$$

$$y'(x) = \sum_{k=0}^{\infty} (2k+s) B_k x^{2k+s-1}, \tag{6.5.4}$$

and

$$y''(x) = \sum_{k=0}^{\infty} (2k+s)(2k+s-1) B_k x^{2k+s-2}, \tag{6.5.5}$$

---

[22] This case is much simpler than for arbitrary n. See Hildebrand, F. B., 1962: *Advanced Calculus for Applications*. Prentice-Hall, Section 4.8.

where we formally assume that we can interchange the order of differentiation and summation. The substitution of Equation 6.5.3 and Equation 6.5.5 into Equation 6.5.1 with $\mu = 1$ yields

$$\sum_{k=0}^{\infty}(2k+s)(2k+s-1)B_k x^{2k+s} + \sum_{k=0}^{\infty}(2k+s)B_k x^{2k+s} + \sum_{k=0}^{\infty} B_k x^{2k+s+2} - n^2 \sum_{k=0}^{\infty} B_k x^{2k+s} = 0, \tag{6.5.6}$$

or

$$\sum_{k=0}^{\infty}[(2k+s)^2 - n^2]B_k x^{2k} + \sum_{k=0}^{\infty} B_k x^{2k+2} = 0. \tag{6.5.7}$$

If we explicitly separate the $k = 0$ term from the other terms in the first summation in Equation 6.5.7,

$$(s^2 - n^2)B_0 + \sum_{m=1}^{\infty}[(2m+s)^2 - n^2]B_m x^{2m} + \sum_{k=0}^{\infty} B_k x^{2k+2} = 0. \tag{6.5.8}$$

We now change the dummy integer in the first summation of Equation 6.5.8 by letting $m = k + 1$ so that

$$(s^2 - n^2)B_0 + \sum_{k=0}^{\infty}\{[(2k+s+2)^2 - n^2]B_{k+1} + B_k\}x^{2k+2} = 0. \tag{6.5.9}$$

Because Equation 6.5.9 must be true for all x, each power of x must vanish identically. This yields $s = \pm n$, and

$$[(2k+s+2)^2 - n^2]B_{k+1} + B_k = 0. \tag{6.5.10}$$

Since the difference of the larger indicial root from the lower root equals the integer $2n$, we are only guaranteed a power series solution of the form given by Equation 6.5.3 for $s = n$. If we use this indicial root and the recurrence formula, Equation 6.5.10, this solution, known as the Bessel function of the first kind of order n and denoted by $J_n(x)$, is

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k (x/2)^{n+2k}}{k!(n+k)!}. \tag{6.5.11}$$

To find the second general solution to Bessel's equation, the one corresponding to $s = -n$, the most economical method[23] is to express it in terms of partial derivatives of $J_n(x)$ with respect to its order n:

$$Y_n(x) = \left[\frac{\partial J_\nu(x)}{\partial \nu} - (-1)^n \frac{\partial J_{-\nu}(x)}{\partial \nu}\right]_{\nu=n}. \tag{6.5.12}$$

Upon substituting the power series representation, Equation 6.5.11, into Equation 6.5.12,

$$\begin{aligned} Y_n(x) = {} & \frac{2}{\pi} J_n(x) \ln(x/2) - \frac{1}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!} \left(\frac{x}{2}\right)^{2k-n} \\ & - \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k (x/2)^{n+2k}}{k!(n+k)!}[\psi(k+1) + \psi(k+n+1)], \end{aligned} \tag{6.5.13}$$

[23] See Watson, G. N., 1966: *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, Section 3.5, for the derivation.

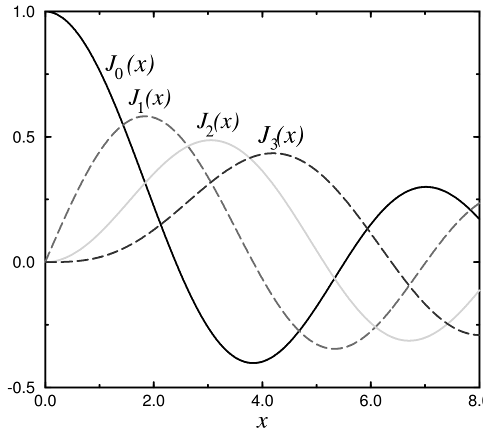


Figure 6.5.1: The first four Bessel functions of the first kind over $0 \leq x \leq 8$.

where

$$\psi(m+1) = -\gamma + 1 + \frac{1}{2} + \cdots + \frac{1}{m}, \tag{6.5.14}$$

$\psi(1) = -\gamma$, and γ is Euler's constant (0.5772157). In the case of $n = 0$, the first sum in Equation 6.5.13 disappears. This function $Y_n(x)$ is Neumann's Bessel function of the second kind of order n. Consequently, the general solution to Equation 6.5.1 is

$$y(x) = AJ_n(\mu x) + BY_n(\mu x). \tag{6.5.15}$$

Figure 6.5.1 illustrates the functions $J_0(x)$, $J_1(x)$, $J_2(x)$, and $J_3(x)$ while Figure 6.5.2 gives $Y_0(x)$, $Y_1(x)$, $Y_2(x)$, and $Y_3(x)$.

An alternative solution to Equation 6.5.1 is

$$y(x) = CH_n^{(1)}(x) + DH_n^{(2)}(x), \tag{6.5.16}$$

where

$$H_n^{(1)}(x) = J_n(x) + iY_n(x), \tag{6.5.17}$$

and

$$H_n^{(2)}(x) = J_n(x) - iY_n(x). \tag{6.5.18}$$

These functions $H_n^{(1)}(x), H_n^{(2)}(x)$ are referred to as Bessel functions of the third kind or *Hankel functions*, after the German mathematician Hermann Hankel (1839–1873). The advantage of Hankel functions over the conventional Bessel function is most clearly seen in their asymptotic expansions:

$$H_n^{(1)}(z) \sim \sqrt{\frac{2}{\pi z}}\, e^{i(z - n\pi/2 - \pi/4)}, \tag{6.5.19}$$

and

$$H_n^{(2)}(z) \sim \sqrt{\frac{2}{\pi z}}\, e^{-i(z - n\pi/2 - \pi/4)}. \tag{6.5.20}$$

for $|z| \to \infty$.

An equation that is very similar to Equation 6.5.1 is

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} - (n^2 + x^2)y = 0. \tag{6.5.21}$$

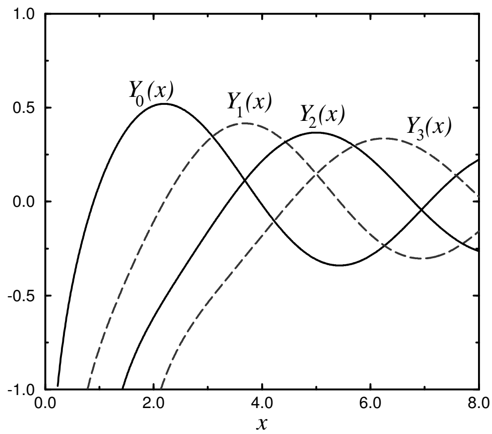


Figure 6.5.2: The first four Bessel functions of the second kind over $0 \leq x \leq 8$.

It arises in the solution of partial differential equations in cylindrical coordinates. If we substitute $ix = t$ (where $i = \sqrt{-1}$) into Equation 6.5.21, it becomes Bessel's equation:

$$t^2 \frac{d^2y}{dt^2} + t\frac{dy}{dt} + (t^2 - n^2)y = 0. \qquad \mathbf{(6.5.22)}$$

Consequently, we may immediately write the solution to Equation 6.5.21 as

$$y(x) = c_1 J_n(ix) + c_2 Y_n(ix), \qquad \mathbf{(6.5.23)}$$

if n is an integer. Traditionally the solution to Equation 6.5.21 has been written

$$y(x) = c_1 I_n(x) + c_2 K_n(x) \qquad \mathbf{(6.5.24)}$$

rather than in terms of $J_n(ix)$ and $Y_n(ix)$, where

$$I_n(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{2k+n}}{k!(k+n)!}, \qquad \mathbf{(6.5.25)}$$

and

$$K_n(x) = \frac{\pi}{2} i^{n+1} \left[J_n(ix) + iY_n(ix) \right]. \qquad \mathbf{(6.5.26)}$$

The function $I_n(x)$ is the modified Bessel function of the first kind, of order n, while $K_n(x)$ is the modified Bessel function of the second kind, of order n. Figure 6.5.3 illustrates $I_0(x)$, $I_1(x)$, $I_2(x)$, and $I_3(x)$ while in Figure 6.5.3 $K_0(x)$, $K_1(x)$, $K_2(x)$, and $K_3(x)$ are graphed. Note that $K_n(x)$ has no real zeros while $I_n(x)$ equals zero only at $x = 0$ for $n \geq 1$.

As our derivation suggests, modified Bessel functions are related to ordinary Bessel functions via complex variables. In particular, $J_n(iz) = i^n I_n(z)$, and $I_n(iz) = i^n J_n(z)$ for z complex.

Although we found solutions to Bessel's equation, Equation 6.5.1, as well as Equation 6.5.21, can we use any of them in an eigenfunction expansion? From Figures 6.5.1–6.5.4 we see that $J_n(x)$ and $I_n(x)$ remain finite at $x = 0$ while $Y_n(x)$ and $K_n(x)$ do not. Furthermore, the products $xJ_n'(x)$ and $xI_n'(x)$ tend to zero at $x = 0$. Thus, both $J_n(x)$ and $I_n(x)$ satisfy the first requirement of an eigenfunction for a Fourier-Bessel expansion.

What about the second condition, that the eigenfunction must satisfy the homogeneous boundary condition, Equation 6.1.2, at $x = L$? From Figure 6.5.3 we see that $I_n(x)$ can

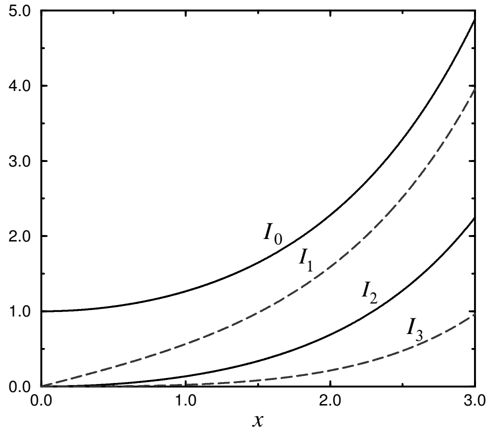


Figure 6.5.3: The first four modified Bessel functions of the first kind over $0 \le x \le 3$.

never satisfy this condition, while from Figure 6.5.1, $J_n(x)$ can. For that reason, we discard $I_n(x)$ from further consideration and continue our analysis only with $J_n(x)$.

Before we can derive the expressions for a Fourier-Bessel expansion, we need to find how $J_n(x)$ is related to $J_{n+1}(x)$ and $J_{n-1}(x)$. Assuming that n is a positive integer, we multiply the series, Equation 6.5.11, by x^n and then differentiate with respect to x. This gives

$$\frac{d}{dx}\left[x^n J_n(x)\right] = \sum_{k=0}^{\infty} \frac{(-1)^k (2n+2k) x^{2n+2k-1}}{2^{n+2k} k!(n+k)!} = x^n \sum_{k=0}^{\infty} \frac{(-1)^k (x/2)^{n-1+2k}}{k!(n-1+k)!} = x^n J_{n-1}(x) \tag{6.5.27}$$

or

$$\boxed{\frac{d}{dx}[x^n J_n(x)] = x^n J_{n-1}(x)} \tag{6.5.28}$$

for $n = 1, 2, 3, \ldots$. Similarly, multiplying Equation 6.5.11 by x^{-n}, we find that

$$\boxed{\frac{d}{dx}\left[x^{-n} J_n(x)\right] = -x^{-n} J_{n+1}(x)} \tag{6.5.29}$$

for $n = 0, 1, 2, 3, \ldots$. If we now carry out the differentiation on Equation 6.5.28 and Equation 6.5.29 and divide by the factors $x^{\pm n}$, we have that

$$J_n'(x) + \frac{n}{x} J_n(x) = J_{n-1}(x), \tag{6.5.30}$$

and

$$J_n'(x) - \frac{n}{x} J_n(x) = -J_{n+1}(x). \tag{6.5.31}$$

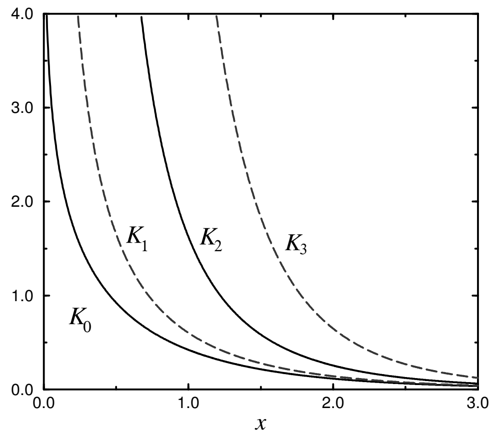


Figure 6.5.4: The first four modified Bessel functions of the second kind over $0 \leq x \leq 3$.

Equation 6.3.30 and Equation 6.3.31 immediately yield the *recurrence relationships*

$$J_{n-1}(x) + J_{n+1}(x) = \frac{2n}{x} J_n(x) \tag{6.5.32}$$

and

$$J_{n-1}(x) - J_{n+1}(x) = 2J_n'(x) \tag{6.5.33}$$

for $n = 1, 2, 3, \ldots$. For $n = 0$, we replace Equation 6.5.33 by $J_0'(x) = -J_1(x)$. Many of the most useful recurrence formulas are summarized in Table 6.5.1 for Bessel functions and in Table 6.5.2 for Hankel functions.

Let us now construct a Fourier-Bessel series. The exact form of the expansion depends upon the boundary condition at $x = L$. There are three possible cases. One of them is $y(L) = 0$ and results in the condition that $J_n(\mu_k L) = 0$. Another condition is $y'(L) = 0$ and gives $J_n'(\mu_k L) = 0$. Finally, if $hy(L) + y'(L) = 0$, then $hJ_n(\mu_k L) + \mu_k J_n'(\mu_k L) = 0$. In all of these cases, the eigenfunction expansion is the same, namely

$$f(x) = \sum_{k=1}^{\infty} A_k J_n(\mu_k x), \tag{6.5.34}$$

where μ_k is the kth positive solution of either $J_n(\mu_k L) = 0$, $J_n'(\mu_k L) = 0$, or $hJ_n(\mu_k L) + \mu_k J_n'(\mu_k L) = 0$.

We now need a mechanism for computing A_k. We begin by multiplying Equation 6.5.34 by $xJ_n(\mu_m x)\,dx$ and integrate from 0 to L. This yields

$$\sum_{k=1}^{\infty} A_k \int_0^L xJ_n(\mu_k x)J(\mu_m x)\,dx = \int_0^L xf(x)J_n(\mu_m x)\,dx. \tag{6.5.35}$$

Table 6.5.1: Some Useful Relationships Involving Bessel Functions of Integer Order

$$J_{n-1}(z) + J_{n+1}(z) = \frac{2n}{z} J_n(z), \qquad n = 1, 2, 3, \ldots$$

$$J_{n-1}(z) - J_{n+1}(z) = 2J'_n(z), \; n = 1, 2, 3, \ldots; \quad J'_0(z) = -J_1(z)$$

$$\frac{d}{dz}\Big[z^n J_n(z)\Big] = z^n J_{n-1}(z), \qquad n = 1, 2, 3, \ldots$$

$$\frac{d}{dz}\Big[z^{-n} J_n(z)\Big] = -z^{-n} J_{n+1}(z), \qquad n = 0, 1, 2, 3, \ldots$$

$$I_{n-1}(z) - I_{n+1}(z) = \frac{2n}{z} I_n(z), \qquad n = 1, 2, 3, \ldots$$

$$I_{n-1}(z) + I_{n+1}(z) = 2I'_n(z), \; n = 1, 2, 3, \ldots; \quad I'_0(z) = I_1(z)$$

$$K_{n-1}(z) - K_{n+1}(z) = -\frac{2n}{z} K_n(z), \qquad n = 1, 2, 3, \ldots$$

$$K_{n-1}(z) + K_{n+1}(z) = -2K'_n(z), \; n = 1, 2, 3, \ldots; \quad K'_0(z) = -K_1(z)$$

$$J_n(ze^{m\pi i}) = e^{nm\pi i} J_n(z)$$

$$I_n(ze^{m\pi i}) = e^{nm\pi i} I_n(z)$$

$$K_n(ze^{m\pi i}) = e^{-mn\pi i} K_n(z) - m\pi i \frac{\cos(mn\pi)}{\cos(n\pi)} I_n(z)$$

$$I_n(z) = e^{-n\pi i/2} J_n(ze^{\pi i/2}), \qquad -\pi < \arg(z) \le \pi/2$$

$$I_n(z) = e^{3n\pi i/2} J_n(ze^{-3\pi i/2}), \qquad \pi/2 < \arg(z) \le \pi$$

From the general orthogonality condition, Equation 6.2.1,

$$\int_0^L x J_n(\mu_k x) J_n(\mu_m x)\, dx = 0, \tag{6.5.36}$$

if $k \neq m$. Equation 6.5.35 then simplifies to

$$A_m \int_0^L x J_n^2(\mu_m x)\, dx = \int_0^L x f(x) J_n(\mu_m x)\, dx, \tag{6.5.37}$$

or

$$\boxed{A_k = \frac{1}{C_k} \int_0^L x f(x) J_n(\mu_k x)\, dx,} \tag{6.5.38}$$

where

$$C_k = \int_0^L x J_n^2(\mu_k x)\, dx, \tag{6.5.39}$$

Table 6.5.2: Some Useful Recurrence Relations for Hankel Functions

$$\frac{d}{dx}\left[x^n H_n^{(p)}(x)\right] = x^n H_{n-1}^{(p)}(x), \; n = 1, 2, \ldots; \; \frac{d}{dx}\left[H_0^{(p)}(x)\right] = -H_1^{(p)}(x)$$

$$\frac{d}{dx}\left[x^{-n} H_n^{(p)}(x)\right] = -x^{-n} H_{n+1}^{(p)}(x), \qquad n = 0, 1, 2, 3, \ldots$$

$$H_{n-1}^{(p)}(x) + H_{n+1}^{(p)}(x) = \frac{2n}{x} H_n^{(p)}(x), \qquad n = 1, 2, 3, \ldots$$

$$H_{n-1}^{(p)}(x) - H_{n+1}^{(p)}(x) = 2\frac{dH_n^{(p)}(x)}{dx}, \qquad n = 1, 2, 3, \ldots$$

and k replaces m in Equation 6.5.37.

The factor C_k depends upon the nature of the boundary conditions at $x = L$. In all cases we start from Bessel's equation

$$[xJ_n'(\mu_k x)]' + \left(\mu_k^2 x - \frac{n^2}{x}\right) J_n(\mu_k x) = 0. \qquad \textbf{(6.5.40)}$$

If we multiply both sides of Equation 6.5.40 by $2xJ_n'(\mu_k x)$, the resulting equation is

$$\left(\mu_k^2 x^2 - n^2\right) \left[J_n^2(\mu_k x)\right]' = -\frac{d}{dx}[xJ_n'(\mu_k x)]^2. \qquad \textbf{(6.5.41)}$$

An integration of Equation 6.5.41 from 0 to L, followed by the subsequent use of integration by parts, results in

$$(\mu_k^2 x^2 - n^2) J_n^2(\mu_k x)\Big|_0^L - 2\mu_k^2 \int_0^L x J_n^2(\mu_k x)\, dx = -\left[xJ_n'(\mu_k x)\right]^2\Big|_0^L. \qquad \textbf{(6.5.42)}$$

Because $J_n(0) = 0$ for $n > 0$, $J_0(0) = 1$ and $xJ_n'(x) = 0$ at $x = 0$, the contribution from the lower limits vanishes. Thus,

$$C_k = \int_0^L x J_n^2(\mu_k x)\, dx = \frac{1}{2\mu_k^2}\left[(\mu_k^2 L^2 - n^2) J_n^2(\mu_k L) + L^2 {J_n'}^2(\mu_k L)\right]. \qquad \textbf{(6.5.43)}$$

Because

$$J_n'(\mu_k x) = \frac{n}{x} J_n(\mu_k x) - \mu_k J_{n+1}(\mu_k x) \qquad \textbf{(6.5.44)}$$

from Equation 6.5.31, C_k becomes

$$\boxed{C_k = \tfrac{1}{2} L^2 J_{n+1}^2(\mu_k L),} \qquad \textbf{(6.5.45)}$$

if $J_n(\mu_k L) = 0$. Otherwise, if $J'_n(\mu_k L) = 0$, then

$$C_k = \frac{\mu_k^2 L^2 - n^2}{2\mu_k^2} J_n^2(\mu_k L). \tag{6.5.46}$$

Finally,

$$C_k = \frac{\mu_k^2 L^2 - n^2 + h^2 L^2}{2\mu_k^2} J_n^2(\mu_k L), \tag{6.5.47}$$

if $\mu_k J'_n(\mu_k L) = -h J_n(\mu_k L)$.

All of the preceding results must be slightly modified when $n = 0$ and the boundary condition is $J'_0(\mu_k L) = 0$ or $\mu_k J_1(\mu_k L) = 0$. This modification results from the additional eigenvalue $\mu_0 = 0$ being present and we must add the extra term A_0 to the expansion. For this case the series reads

$$f(x) = A_0 + \sum_{k=1}^{\infty} A_k J_0(\mu_k x), \tag{6.5.48}$$

where the equation for finding A_0 is

$$A_0 = \frac{2}{L^2} \int_0^L f(x)\, x\, dx, \tag{6.5.49}$$

and Equation 6.5.38 and Equation 6.5.46 with $n = 0$ give the remaining coefficients.

• Example 6.5.1

Starting with Bessel's equation, we show that the solution to

$$y'' + \frac{1-2a}{x} y' + \left(b^2 c^2 x^{2c-2} + \frac{a^2 - n^2 c^2}{x^2} \right) y = 0 \tag{6.5.50}$$

is

$$y(x) = A x^a J_n(b x^c) + B x^a Y_n(b x^c), \tag{6.5.51}$$

provided that $bx^c > 0$ so that $Y_n(bx^c)$ exists.

The general solution to

$$\xi^2 \frac{d^2\eta}{d\xi^2} + \xi \frac{d\eta}{d\xi} + (\xi^2 - n^2)\eta = 0 \tag{6.5.52}$$

is

$$\eta = A J_n(\xi) + B Y_n(\xi). \tag{6.5.53}$$

If we now let $\eta = y(x)/x^a$ and $\xi = bx^c$, then

$$\frac{d}{d\xi} = \frac{dx}{d\xi} \frac{d}{dx} = \frac{x^{1-c}}{bc} \frac{d}{dx}, \tag{6.5.54}$$

$$\frac{d^2}{d\xi^2} = \frac{x^{2-2c}}{b^2c^2}\frac{d^2}{dx^2} - \frac{(c-1)x^{1-2c}}{b^2c^2}\frac{d}{dx}, \tag{6.5.55}$$

$$\frac{d}{dx}\left(\frac{y}{x^a}\right) = \frac{1}{x^a}\frac{dy}{dx} - \frac{a}{x^{a+1}}y, \tag{6.5.56}$$

and

$$\frac{d^2}{dx^2}\left(\frac{y}{x^a}\right) = \frac{1}{x^a}\frac{d^2y}{dx^2} - \frac{2a}{x^{a+1}}\frac{dy}{dx} + \frac{a(1+a)}{x^{a+2}}y. \tag{6.5.57}$$

Substituting Equation 6.5.54 and Equation 6.5.57 into Equation 6.5.52 and simplifying yields the desired result. □

- **Example 6.5.2**

Let us find[24] the general solution to the nonhomogeneous differential equation

$$\frac{d^2y}{dr^2} + \frac{1}{r}\frac{dy}{dr} - k^2y = -S(r), \tag{6.5.58}$$

where k is a real parameter.

The homogeneous solution is

$$y_H(r) = C_1I_0(kr) + C_2K_0(kr). \tag{6.5.59}$$

Using variation of parameters, we assume that the particular solution can be written

$$y_p(r) = A(r)I_0(kr) + B(r)K_0(kr), \tag{6.5.60}$$

where

$$A'(r) = \begin{vmatrix} 0 & K_0(kr) \\ -S(r) & kK_0'(kr) \end{vmatrix} \Big/ \begin{vmatrix} I_0(kr) & K_0(kr) \\ kI_0'(kr) & kK_0'(kr) \end{vmatrix}, \tag{6.5.61}$$

and

$$B'(r) = \begin{vmatrix} I_0(kr) & 0 \\ kI_0'(kr) & -S(r) \end{vmatrix} \Big/ \begin{vmatrix} I_0(kr) & K_0(kr) \\ kI_0'(kr) & kK_0'(kr) \end{vmatrix}. \tag{6.5.62}$$

Expanding the determinants, we find

$$A'(r) = S(r)K_0(kr)/\left\{k\left[I_0(kr)K_0'(kr) - I_0'(kr)K_0(kr)\right]\right\} \tag{6.5.63}$$

and

$$B'(r) = -S(r)I_0(kr)/\left\{k\left[I_0(kr)K_0'(kr) - I_0'(kr)K_0(kr)\right]\right\}. \tag{6.5.64}$$

Evaluating the Wronskian[25] for modified Bessel functions,

$$I_0(z)K_0'(z) - I_0'(z)K_0(z) = -1/z, \tag{6.5.65}$$

$$A'(r) = -rS(r)K_0(kr) \qquad \text{and} \qquad B'(r) = rS(r)I_0(kr). \tag{6.5.66}$$

---

[24] See Hassan, M. H. A., 1988: Ion distribution functions during ion cyclotron resonance heating at the fundamental frequency. *Phys. Fluids*, **31**, 596–599.

[25] Watson, op. cit., p. 80, Formula 19.

Integrating Equation 6.5.66,

$$A(r) = -\int^r xS(x)K_0(kx)\,dx, \qquad \text{and} \qquad B(r) = \int^r xS(x)I_0(kx)\,dx. \qquad \textbf{(6.5.67)}$$

Consequently, the general solution is the sum of the particular and homogeneous solution,

$$y(r) = C_1 I_0(kr) + C_2 K_0(kr) - I_0(kr)\int^r xS(x)K_0(kx)\,dx + K_0(kr)\int^r xS(x)I_0(kx)\,dx. \qquad \textbf{(6.5.68)}$$

□

• **Example 6.5.3**

Let us show that

$$x^2 J_n''(x) = (n^2 - n - x^2)J_n(x) + xJ_{n+1}(x). \qquad \textbf{(6.5.69)}$$

From Equation 6.5.31,

$$J_n'(x) = \frac{n}{x}J_n(x) - J_{n+1}(x), \qquad \textbf{(6.5.70)}$$

$$J_n''(x) = -\frac{n}{x^2}J_n(x) + \frac{n}{x}J_n'(x) - J_{n+1}'(x), \qquad \textbf{(6.5.71)}$$

and

$$J_n''(x) = -\frac{n}{x^2}J_n(x) + \frac{n}{x}\left[\frac{n}{x}J_n(x) - J_{n+1}(x)\right] - \left[J_n(x) - \frac{n+1}{x}J_{n+1}(x)\right] \qquad \textbf{(6.5.72)}$$

after using Equation 6.5.30 and Equation 6.5.31. Simplifying,

$$J_n''(x) = \left(\frac{n^2-n}{x^2} - 1\right)J_n(x) + \frac{J_{n+1}(x)}{x}. \qquad \textbf{(6.5.73)}$$

After multiplying Equation 6.5.73 by x^2, we obtain Equation 6.5.69. □

• **Example 6.5.4**

Let us show that

$$\int_0^a x^5 J_2(x)\,dx = a^5 J_3(a) - 2a^4 J_4(a). \qquad \textbf{(6.5.74)}$$

We begin by integrating Equation 6.5.74 by parts. If $u = x^2$, and $dv = x^3 J_2(x)\,dx$, then

$$\int_0^a x^5 J_2(x)\,dx = x^5 J_3(x)\big|_0^a - 2\int_0^a x^4 J_3(x)\,dx, \qquad \textbf{(6.5.75)}$$

because $d[x^3 J_3(x)]/dx = x^2 J_2(x)$ by Equation 6.5.28. Finally,

$$\int_0^a x^5 J_2(x)\,dx = a^5 J_3(a) - 2x^4 J_4(x)\big|_0^a = a^5 J_3(a) - 2a^4 J_4(a), \qquad \textbf{(6.5.76)}$$

since $x^4 J_3(x) = d[x^4 J_4(x)]/dx$ by Equation 6.5.28. □

• **Example 6.5.5**

Let us expand $f(x) = x$, $0 < x < 1$, in the series

$$f(x) = \sum_{k=1}^{\infty} A_k J_1(\mu_k x), \tag{6.5.77}$$

where μ_k denotes the kth zero of $J_1(\mu)$. From Equation 6.5.38 and Equation 6.5.46,

$$A_k = \frac{2}{J_2^2(\mu_k)} \int_0^1 x^2 J_1(\mu_k x)\, dx. \tag{6.5.78}$$

However, from Equation 6.5.28,

$$\frac{d}{dx}\left[x^2 J_2(x)\right] = x^2 J_1(x), \tag{6.5.79}$$

if $n = 2$. Therefore, Equation 6.5.78 becomes

$$A_k = \left.\frac{2x^2 J_2(x)}{\mu_k^3 J_2^2(\mu_k)}\right|_0^{\mu_k} = \frac{2}{\mu_k J_2(\mu_k)}, \tag{6.5.80}$$

and the resulting expansion is

$$x = 2\sum_{k=1}^{\infty} \frac{J_1(\mu_k x)}{\mu_k J_2(\mu_k)}, \qquad 0 \le x < 1. \tag{6.5.81}$$

Figure 6.5.5 shows the Fourier-Bessel expansion of $f(x) = x$ in truncated form when we only include one, two, three, and four terms. It was created using the MATLAB script

```
clear;
x = [0:0.01:1]; % create x points in plot
f = x; % construct function f(x)
% initialize Fourier-Bessel series
fbessel = zeros(size(x));
% read in the first four zeros of J_1(mu) = 0
mu(1) =  3.83171; mu(2) =  7.01559;
mu(3) = 10.17347; mu(4) = 13.32369;
clf % clear any figures
for n = 1:4
% Fourier coefficient
  factor = 2 / (mu(n) * besselj(2,mu(n)));
% compute Fourier-Bessel series
  fbessel = fbessel + factor * besselj(1,mu(n)*x);
% create plot of truncated Fourier-Bessel series
%     with n terms
  subplot(2,2,n), plot(x,fbessel,x,f,'--')
  axis([0 1 -0.25 1.25])
  if n == 1 legend('1 term','f(x)'); legend boxoff;
    else legend([num2str(n) ' terms'],'f(x)'); legend boxoff;
  end
```

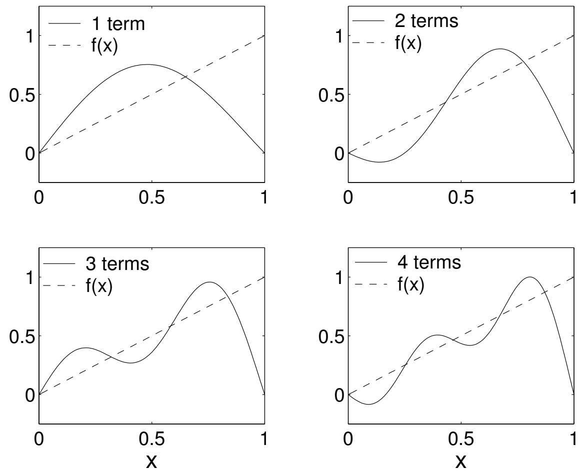


Figure 6.5.5: The Fourier-Bessel series representation, Equation 6.5.81, for $f(x) = x$, $0 < x < 1$, when we truncate the series so that it includes only the first, first two, first three, and first four terms.

```
  if n > 2 xlabel('x','Fontsize',20); end
end
```

□

• **Example 6.5.6**

Let us expand the function $f(x) = x^2$, $0 < x < 1$, in the series

$$f(x) = \sum_{k=1}^{\infty} A_k J_0(\mu_k x), \qquad (6.5.82)$$

where μ_k denotes the kth positive zero of $J_0(\mu)$. From Equation 6.5.38 and Equation 6.5.46,

$$A_k = \frac{2}{J_1^2(\mu_k)} \int_0^1 x^3 J_0(\mu_k x)\, dx. \qquad (6.5.83)$$

If we let $t = \mu_k x$, the integration, Equation 6.5.83, becomes

$$A_k = \frac{2}{\mu_k^4 J_1^2(\mu_k)} \int_0^{\mu_k} t^3 J_0(t)\, dt. \qquad (6.5.84)$$

We now let $u = t^2$ and $dv = tJ_0(t)\, dt$ so that integration by parts results in

$$A_k = \frac{2}{\mu_k^4 J_1^2(\mu_k)} \left[t^3 J_1(t) \Big|_0^{\mu_k} - 2 \int_0^{\mu_k} t^2 J_1(t)\, dt \right] = \frac{2}{\mu_k^4 J_1^2(\mu_k)} \left[\mu_k^3 J_1(\mu_k) - 2 \int_0^{\mu_k} t^2 J_1(t)\, dt \right], \qquad (6.5.85)$$

because $v = tJ_1(t)$ from Equation 6.5.28. If we integrate by parts once more, we find that

$$A_k = \frac{2}{\mu_k^4 J_1^2(\mu_k)} \left[\mu_k^3 J_1(\mu_k) - 2\mu_k^2 J_2(\mu_k) \right] = \frac{2}{J_1^2(\mu_k)} \left[\frac{J_1(\mu_k)}{\mu_k} - \frac{2J_2(\mu_k)}{\mu_k^2} \right]. \qquad (6.5.86)$$

However, from Equation 6.5.32 with $n = 1$,

$$J_1(\mu_k) = \tfrac{1}{2}\mu_k \left[J_2(\mu_k) + J_0(\mu_k) \right], \qquad (6.5.87)$$

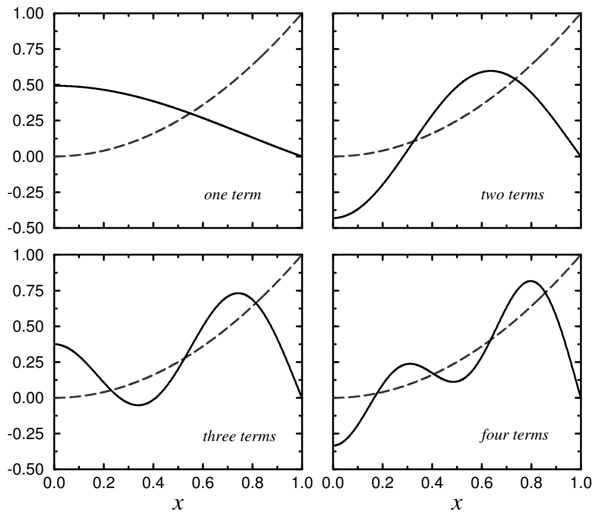


Figure 6.5.6: The Fourier-Bessel series representation, Equation 6.5.90, for $f(x) = x^2$, $0 < x < 1$, when we truncate the series so that it includes only the first, first two, first three, and first four terms.

or

$$J_2(\mu_k) = \frac{2J_1(\mu_k)}{\mu_k}, \tag{6.5.88}$$

because $J_0(\mu_k) = 0$. Therefore,

$$A_k = \frac{2(\mu_k^2 - 4)J_1(\mu_k)}{\mu_k^3 J_1^2(\mu_k)}, \tag{6.5.89}$$

and

$$x^2 = 2\sum_{k=1}^{\infty} \frac{(\mu_k^2 - 4)J_0(\mu_k x)}{\mu_k^3 J_1(\mu_k)}, \qquad 0 < x < 1. \tag{6.5.90}$$

Figure 6.5.6 shows the representation of x^2 by the Fourier-Bessel series, Equation 6.5.90 when we truncate it so that it includes only one, two, three, or four terms. As we add each additional term in the orthogonal expansion, the expansion fits $f(x)$ better in the "least squares" sense of Equation 6.3.5.

Problems

1. Show from the series solution that

$$\frac{d}{dx}\left[J_0(kx)\right] = -kJ_1(kx).$$

From the recurrence formulas, show the following relations:

2. $2J_0''(x) = J_2(x) - J_0(x)$

3. $J_2(x) = J_0''(x) - J_0'(x)/x$

4.

$$J_0'''(x) = \frac{J_0(x)}{x} + \left(\frac{2}{x^2} - 1\right)J_0'(x)$$

5.

$$\frac{J_2(x)}{J_1(x)} = \frac{1}{x} - \frac{J_0''(x)}{J_0'(x)} = \frac{2}{x} - \frac{J_0(x)}{J_1(x)} = \frac{2}{x} + \frac{J_0(x)}{J_0'(x)}$$

6.

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

7.

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

8.

$$J_3(x) = \left(\frac{8}{x^2} - 1\right) J_1(x) - \frac{4}{x} J_0(x)$$

9.

$$4J_n''(x) = J_{n-2}(x) - 2J_n(x) + J_{n+2}(x)$$

10. Show that the maximum and minimum values of $J_n(x)$ occur when

$$x = \frac{nJ_n(x)}{J_{n+1}(x)}, \quad x = \frac{nJ_n(x)}{J_{n-1}(x)}, \quad \text{and} \quad J_{n-1}(x) = J_{n+1}(x).$$

Show that

11.

$$\frac{d}{dx}\left[x^2 J_3(2x)\right] = -xJ_3(2x) + 2x^2 J_2(2x)$$

12.

$$\frac{d}{dx}\left[xJ_0(x^2)\right] = J_0(x^2) - 2x^2 J_1(x^2)$$

13.

$$\int x^3 J_2(3x)\, dx = \tfrac{1}{3}x^3 J_3(3x) + C$$

14.

$$\int x^{-2} J_3(2x)\, dx = -\tfrac{1}{2}x^{-2} J_2(2x) + C$$

15.

$$\int x \ln(x) J_0(x)\, dx = J_0(x) + x\ln(x)J_1(x) + C$$

16.

$$\int_0^a xJ_0(kx)\, dx = \frac{a^2 J_1(ka)}{ka}$$

17.

$$\int_0^1 x(1-x^2)J_0(kx)\,dx = \frac{4}{k^3}J_1(k) - \frac{2}{k^2}J_0(k)$$

18.

$$\int_0^1 x^3 J_0(kx)\,dx = \frac{k^2-4}{k^3}J_1(k) + \frac{2}{k^2}J_0(k)$$

19. Show that

$$1 = 2\sum_{k=1}^{\infty} \frac{J_0(\mu_k x)}{\mu_k J_1(\mu_k)}, \qquad 0 \le x < 1,$$

where μ_k is the kth positive root of $J_0(\mu) = 0$. Then use MATLAB to illustrate various partial sums of the Fourier-Bessel series.

20. Show that

$$\frac{1-x^2}{8} = \sum_{k=1}^{\infty} \frac{J_0(\mu_k x)}{\mu_k^3 J_1(\mu_k)}, \qquad 0 \le x \le 1,$$

where μ_k is the kth positive root of $J_0(\mu) = 0$. Then use MATLAB to illustrate various partial sums of the Fourier-Bessel series.

21. Show that

$$4x - x^3 = -16\sum_{k=1}^{\infty} \frac{J_1(\mu_k x)}{\mu_k^3 J_0(2\mu_k)}, \qquad 0 \le x \le 2,$$

where μ_k is the kth positive root of $J_1(2\mu) = 0$. Then use MATLAB to illustrate various partial sums of the Fourier-Bessel series.

22. Show that

$$x^3 = 2\sum_{k=1}^{\infty} \frac{(\mu_k^2 - 8)J_1(\mu_k x)}{\mu_k^3 J_2(\mu_k)}, \qquad 0 \le x \le 1,$$

where μ_k is the kth positive root of $J_1(\mu) = 0$. Then use MATLAB to illustrate various partial sums of the Fourier-Bessel series.

23. Show that

$$x = 2\sum_{k=1}^{\infty} \frac{\mu_k J_2(\mu_k) J_1(\mu_k x)}{(\mu_k^2 - 1)J_1^2(\mu_k)}, \qquad 0 \le x \le 1,$$

where μ_k is the kth positive root of $J_1'(\mu) = 0$. Then use MATLAB to illustrate various partial sums of the Fourier-Bessel series.

24. Show that

$$1 - x^4 = 32\sum_{k=1}^{\infty} \frac{(\mu_k^2 - 4)J_0(\mu_k x)}{\mu_k^5 J_1(\mu_k)}, \qquad 0 \le x \le 1,$$

where μ_k is the kth positive root of $J_0(\mu) = 0$. Then use MATLAB to illustrate various partial sums of the Fourier-Bessel series.

25. Show that

$$1 = 2\alpha L \sum_{k=1}^{\infty} \frac{J_0(\mu_k x/L)}{(\mu_k^2 + \alpha^2 L^2) J_0(\mu_k)}, \qquad 0 \le x \le L,$$

where μ_k is the kth positive root of $\mu J_1(\mu) = \alpha L J_0(\mu)$. Then use MATLAB to illustrate various partial sums of the Fourier-Bessel series.

26. Using the relationship[26]

$$\int_0^a J_\nu(\alpha r) J_\nu(\beta r)\, r\, dr = \frac{a\beta J_\nu(\alpha a) J_\nu'(\beta a) - a\alpha J_\nu(\beta a) J_\nu'(\alpha a)}{\alpha^2 - \beta^2},$$

show that

$$\frac{J_0(bx) - J_0(ba)}{J_0(ba)} = \frac{2b^2}{a} \sum_{k=1}^{\infty} \frac{J_0(\mu_k x)}{\mu_k(\mu_k^2 - b^2) J_1(\mu_k a)}, \qquad 0 \le x \le a,$$

where μ_k is the kth positive root of $J_0(\mu a) = 0$ and b is a constant.

27. Given the definite integral[27]

$$\int_0^1 \frac{x\, J_0(bx)}{\sqrt{1 - x^2}}\, dx = \frac{\sin(b)}{b}, \qquad 0 < b,$$

show that

$$\frac{H(t - x)}{\sqrt{t^2 - x^2}} = 2 \sum_{k=1}^{\infty} \frac{\sin(\mu_k t) J_0(\mu_k x)}{\mu_k J_1^2(\mu_k)}, \qquad 0 < x < 1, \quad 0 < t \le 1,$$

where μ_k is the kth positive root of $J_0(\mu) = 0$ and $H(\cdot)$ is Heaviside's step function.

28. Using the same definite integral from the previous problem, show[28] that

$$\frac{H(a - x)}{\sqrt{a^2 - x^2}} = \frac{2}{b} \sum_{n=1}^{\infty} \frac{\sin(\mu_n a/b) J_0(\mu_n x/b)}{\mu_n J_0^2(\mu_n)}, \qquad 0 \le x < b,$$

where $a < b$, μ_n is the nth positive root of $J_0'(\mu) = -J_1(\mu) = 0$, and $H(\cdot)$ is Heaviside's step function.

29. Given the definite integral[29]

$$\int_0^a \cos(cx)\; J_0\left(b\sqrt{a^2 - x^2}\right)\, dx = \frac{\sin\left(a\sqrt{b^2 + c^2}\right)}{\sqrt{b^2 + c^2}}, \qquad 0 < b,$$

[26] Watson, op. cit., Section 5.11, Equation 8.

[27] Gradshteyn, I. S., and I. M. Ryzhik, 1965: *Table of Integrals, Series, and Products*. Academic Press, Section 6.567, Formula 1 with $\nu = 0$ and $\mu = -1/2$.

[28] See Wei, X. X., 2000: Finite solid circular cylinders subjected to arbitrary surface load. Part II–Application to double-punch test. *Int. J. Solids Struct.*, **37**, 5733–5744.

[29] Gradshteyn and Ryzhik, op. cit., Section 6.677, Formula 6.

show that

$$\frac{\cosh\left(b\sqrt{t^2-x^2}\,\right)}{\sqrt{t^2-x^2}}H(t-x) = \frac{2}{a^2}\sum_{k=1}^{\infty}\frac{\sin\left(t\sqrt{\mu_k^2-b^2}\,\right)J_0(\mu_k x)}{\sqrt{\mu_k^2-b^2}\,J_1^2(\mu_k a)},$$

where $0 < x < a$, μ_k is the kth positive root of $J_0(\mu a) = 0$, $H(\cdot)$ is Heaviside's step function, and b is a constant.

30. Using the integral definition of the Bessel function[30] for $J_1(z)$:

$$J_1(z) = \frac{2}{\pi}\int_0^1 \frac{t\,\sin(zt)}{\sqrt{1-t^2}}\,dt, \qquad 0 < z,$$

show that

$$\frac{x}{t\sqrt{t^2-x^2}}H(t-x) = \frac{\pi}{L}\sum_{n=1}^{\infty} J_1\left(\frac{n\pi t}{L}\right)\sin\left(\frac{n\pi x}{L}\right), \qquad 0 \le x < L,$$

where $H(\cdot)$ is Heaviside's step function. Hint: Treat this as a Fourier half-range sine expansion.

31. Show that

$$\delta(x-b) = \frac{2b}{a^2}\sum_{k=1}^{\infty}\frac{J_0(\mu_k b/a)J_0(\mu_k x/a)}{J_1^2(\mu_k)}, \qquad 0 \le x, b < a,$$

where μ_k is the kth positive root of $J_0(\mu) = 0$ and $\delta(\cdot)$ is the Dirac delta function.

32. Show that

$$\frac{\delta(x)}{2\pi x} = \frac{1}{\pi a^2}\sum_{k=1}^{\infty}\frac{J_0(\mu_k x/a)}{J_1^2(\mu_k)}, \qquad 0 \le x < a,$$

where μ_k is the kth positive root of $J_0(\mu) = 0$ and $\delta(\cdot)$ is the Dirac delta function.

6.6 FINITE ELEMENT METHOD

In Section 1.7 we showed how to solve ordinary differential equations using finite differences. Here we introduce a popular alternative, the finite element method, and will use it to solve the Sturm-Liouville problem. One advantage of this approach is that we can focus on the details of the numerical scheme.

The finite element method breaks the global solution domain into a number of simply shaped subdomains, called *elements*. The global solution is then constructed by assembling the results from all of the elements. A particular strength of this method is that the elements do not have to be the same size; this allows us to have more resolution in regions where the solution is rapidly changing and fewer elements where the solution changes slowly. Overall the solution of the ordinary differential equation is given by a succession of piecewise continuous functions.

---

[30] Gradshteyn and Ryzhik, Ibid., Section 3.753, Formula 5.

Consider the Sturm-Liouville problem

$$-\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)y(x) - \lambda r(x)y(x) = 0, \qquad a < x < b, \tag{6.6.1}$$

with

$$y(a) = 0, \qquad \text{or} \qquad p(a)y'(a) = 0, \tag{6.6.2}$$

and

$$y(b) = 0, \qquad \text{or} \qquad p(b)y'(b) = 0. \tag{6.6.3}$$

Our formulation of the finite element approximation to the exact solution is called the *Galerkin weighted residual approach.* This is not the only possible way of formulating the finite element equations but it is similar to the eigenfunction expansions that we highlighted in this chapter. Our approach consists of two steps: First we assume that $y(x)$ can be expressed over a particular element by

$$y(x) = \sum_{j=1}^{J} y_j \varphi_j(x), \tag{6.6.4}$$

where $\varphi_j(x)$ is the jth *approximation* or *shape function*, and J is the total number of elements.

Let us define the residue

$$R(x) = -\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)y(x) - \lambda r(x)y(x). \tag{6.6.5}$$

We now require that for each element along the segment $\Omega_e = (x_n, x_{n+1})$,

$$\int_{\Omega_e} R(x)\varphi_i(x)\,dx = 0, \qquad i = 1, 2, 3, \ldots, J. \tag{6.6.6}$$

The points x_n and x_{n+1} are known as *nodes*. Substituting Equation 6.6.5 into Equation 6.6.6,

$$\int_{\Omega_e} \left\{-\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)y(x) - \lambda r(x)y(x)\right\}\varphi_i(x)\,dx = 0. \tag{6.6.7}$$

Because

$$-\int_{x_n}^{x_{n+1}} \frac{d}{dx}\left[p(x)\frac{dy}{dx}\right]\varphi_i(x)\,dx = -p(x)\frac{dy}{dx}\varphi_i(x)\bigg|_{x_n}^{x_{n+1}} + \int_{x_n}^{x_{n+1}} p(x)\frac{dy}{dx}\frac{d\varphi_i(x)}{dx}\,dx, \tag{6.6.8}$$

then

$$\int_{x_n}^{x_{n+1}} \left[p(x)\frac{dy}{dx}\frac{d\varphi_i(x)}{dx} + q(x)y(x)\varphi_i(x) - \lambda r(x)y(x)\phi_i(x)\right] dx = p(x)\frac{dy}{dx}\varphi_i(x)\bigg|_{x_n}^{x_{n+1}}. \tag{6.6.9}$$

Upon using Equation 6.6.4 to eliminate $y(x)$ and reversing the order of summation and integration, our second step in the finite element method involves solving for y_j via

$$\begin{aligned}\sum_{j=1}^{J}\biggl\{\int_{x_n}^{x_{n+1}} p(x)\frac{d\varphi_i(x)}{dx}\frac{d\varphi_j(x)}{dx}\,dx + \int_{x_n}^{x_{n+1}} q(x)\varphi_i(x)\varphi_j(x)\,dx \\ - \lambda\int_{x_n}^{x_{n+1}} r(x)\varphi_i(x)\varphi_j(x)\,dx\biggr\}y_j = p(x)\frac{dy}{dx}\varphi_i(x)\bigg|_{x_n}^{x_{n+1}},\end{aligned} \tag{6.6.10}$$

or using matrix notation

$$K\mathbf{y} - \lambda M\mathbf{y} = \mathbf{b}, \tag{6.6.11}$$

where

$$K = \begin{pmatrix} K_{11} & K_{12} & \dots & K_{1J} \\ K_{21} & K_{22} & \dots & K_{2J} \\ \vdots & \vdots & \vdots & \vdots \\ K_{J1} & K_{J2} & \dots & K_{JJ} \end{pmatrix}, \quad M = \begin{pmatrix} M_{11} & M_{12} & \dots & M_{1J} \\ M_{21} & M_{22} & \dots & M_{2J} \\ \vdots & \vdots & \vdots & \vdots \\ M_{J1} & M_{J2} & \dots & M_{JJ} \end{pmatrix} \tag{6.6.12}$$

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_J \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{pmatrix}, \tag{6.6.13}$$

$$b_i = p(x_{n+1})y'(x_{n+1})\varphi_i(x_{n+1}) - p(x_n)y'(x_n)\varphi_i(x_n), \tag{6.6.14}$$

$$K_{ij} = \int_{x_n}^{x_{n+1}} p(x)\frac{d\varphi_i(x)}{dx}\frac{d\varphi_j(x)}{dx}\,dx + \int_{x_n}^{x_{n+1}} q(x)\varphi_i(x)\varphi_j(x)\,dx, \tag{6.6.15}$$

and

$$M_{ij} = \int_{x_n}^{x_{n+1}} r(x)\varphi_i(x)\varphi_j(x)\,dx. \tag{6.6.16}$$

Why do we prefer to use Equation 6.6.10 rather than Equation 6.6.7? There are two reasons. First, it offers a convenient method for introducing the specified boundary conditions, Equation 6.6.2 and Equation 6.6.3. Second, it has lowered the highest-order derivatives from a second to a first derivative. This yields the significant benefit that $\varphi_i(x)$ must only be continuous but not necessarily a continuous slope at the nodes.

An important question is how we will evaluate the integrals in Equation 6.6.15 and Equation 6.6.16. Because $p(x)$, $q(x)$, and $r(x)$ are known, we could substitute these quantities along with $\varphi_i(x)$ and $\varphi_j(x)$ into Equation 6.6.15 and Equation 6.6.16 and perform the integration, presumably numerically. In a similar vein, we could develop curve fits for $p(x)$, $q(x)$, and $r(x)$ and again perform the integrations. However, we simply use their values at the midpoint between the nodes, $\overline{x}_n = (x_{n+1} + x_n)/2$, because $p(x)$, $q(x)$ and $r(x)$ usually vary slowly over the interval (x_n, x_{n+1}).

At this point we will specify J. The simplest case is $J = 2$ and we have the linear element:

$$\varphi_1(x) = \frac{x - x_1}{x_2 - x_1} \qquad \text{and} \qquad \varphi_2(x) = \frac{x_2 - x}{x_2 - x_1}, \tag{6.6.17}$$

where x_1 and x_2 are *local* nodal points located at the end of the element. It directly follows that

$$\frac{dy}{dx} = \frac{d\varphi_1}{dx}y_1 + \frac{d\varphi_2}{dx}y_2 = \frac{y_2 - y_1}{x_2 - x_1}. \tag{6.6.18}$$

In other words, dy/dx equals the slope of the straight line connecting the nodes. Similarly,

$$\int_{x_1}^{x_2} y(x)\,dx = \tfrac{1}{2}\left(y_2 + y_1\right)\left(x_2 - x_1\right) \tag{6.6.19}$$

and we simply have the trapezoidal rule.

Table 6.6.1: The System Topology for 4 Finite-Element Segmentations When a Linear Interpolation Is Used

	Node Numbers	
Element	Local	Global
1	1	1
	2	2
2	1	2
	2	3
3	1	3
	2	4
4	1	4
	2	5

Substituting $\varphi_1(x)$ and $\varphi_2(x)$ into Equation 6.6.15 and Equation 6.6.16 and carrying out the integration, we obtain

$$K_{11} = \frac{p(x_c)}{L} + \frac{q(x_c)L}{3}, \quad K_{12} = -\frac{p(x_c)}{L} + \frac{q(x_c)L}{6}, \quad K_{21} = K_{12}, \quad \text{and} \quad K_{22} = K_{11}, \tag{6.6.20}$$

with $L = x_2 - x_1$ and $x_c = (x_1 + x_2)/2$. Similarly,

$$M_{11} = \frac{r(x_c)L}{3} = M_{22}, \quad \text{and} \quad M_{12} = \frac{r(x_c)L}{6} = M_{21}. \tag{6.6.21}$$

Finally, because $\varphi_1(x_1) = 0$, $\varphi_1(x_2) = 1$, $\varphi_2(x_1) = 1$, and $\varphi_2(x_2) = 0$, $b_1 = -p(x_1)y'(x_1)$ and $b_2 = p(x_2)y_2(x_2)$.

Having obtained the finite element representation for nodes 1 and 2, we would like to extend these results to an arbitrary number of additional nodes. This is done by setting up a look-up table that relates the global nodal points to the local ones. For example, suppose we would like 5 nodes between a and b with $x = x_1$, x_2, x_3, x_4, and x_5. Then Table 6.6.1 illustrates our look-up table.

Having developed the spatial layout, we are now ready to assemble the matrix for the entire interval (a, b). For clarity we will give the intermediate steps. Taking the first element into account,

$$K = \begin{pmatrix} K_{11}^{(1)} & K_{12}^{(1)} & 0 & 0 & 0 \\ K_{21}^{(1)} & K_{22}^{(1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad M = \begin{pmatrix} M_{11}^{(1)} & M_{12}^{(1)} & 0 & 0 & 0 \\ M_{21}^{(1)} & M_{22}^{(1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{6.6.22}$$

$$\mathbf{b} = \begin{pmatrix} -p(x_1)y'(x_1) \\ p(x_2)y'(x_2) \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{6.6.23}$$

Here we have added a subscript (1) to K_{ij} and M_{ij} to denote that value for the first element should be used in computing $p(x_c)$, $q(x_c)$, $r(x_c)$, and L. Consequently, when we introduce the second element, K, M, and $\mathbf{y}$ become

$$K = \begin{pmatrix} K_{11}^{(1)} & K_{12}^{(1)} & 0 & 0 & 0 \\ K_{21}^{(1)} & K_{22}^{(1)} + K_{11}^{(2)} & K_{12}^{(2)} & 0 & 0 \\ 0 & K_{21}^{(2)} & K_{22}^{(2)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \textbf{(6.6.24)}$$

$$M = \begin{pmatrix} M_{11}^{(1)} & M_{12}^{(1)} & 0 & 0 & 0 \\ M_{21}^{(1)} & M_{22}^{(1)} + M_{11}^{(2)} & M_{12}^{(2)} & 0 & 0 \\ 0 & M_{21}^{(2)} & M_{22}^{(2)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \textbf{(6.6.25)}$$

$$\mathbf{b} = \begin{pmatrix} -p(x_1)y'(x_1) \\ 0 \\ p(x_3)y'(x_3) \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ 0 \\ 0 \end{pmatrix}. \quad \textbf{(6.6.26)}$$

Note that the numbering for y_i corresponds to ith *global* node number. Continuing with this process of adding additional elements to the system matrix, we finally have

$$K = \begin{pmatrix} K_{11}^{(1)} & K_{12}^{(1)} & 0 & 0 & 0 \\ K_{21}^{(1)} & K_{22}^{(1)} + K_{11}^{(2)} & K_{12}^{(2)} & 0 & 0 \\ 0 & K_{21}^{(2)} & K_{22}^{(2)} + K_{11}^{(3)} & K_{12}^{(3)} & 0 \\ 0 & 0 & K_{21}^{(3)} & K_{22}^{(3)} + K_{11}^{(4)} & K_{12}^{(4)} \\ 0 & 0 & 0 & K_{21}^{(4)} & K_{22}^{(4)} \end{pmatrix}, \quad \textbf{(6.6.27)}$$

$$M = \begin{pmatrix} M_{11}^{(1)} & M_{12}^{(1)} & 0 & 0 & 0 \\ M_{21}^{(1)} & M_{22}^{(1)} + M_{11}^{(2)} & M_{12}^{(2)} & 0 & 0 \\ 0 & M_{21}^{(2)} & M_{22}^{(2)} + M_{11}^{(3)} & M_{12}^{(3)} & 0 \\ 0 & 0 & M_{21}^{(3)} & M_{22}^{(3)} + M_{11}^{(4)} & M_{12}^{(4)} \\ 0 & 0 & 0 & M_{21}^{(4)} & M_{22}^{(4)} \end{pmatrix}, \quad \textbf{(6.6.28)}$$

where

$$\mathbf{b} = \begin{pmatrix} -p(x_1)y'(x_1) \\ 0 \\ 0 \\ 0 \\ p(x_5)y'(x_5) \end{pmatrix}, \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}. \quad \textbf{(6.6.29)}$$

Let us examine the $\mathbf{b}$ vector more closely. In the final form of the finite element formulation, $\mathbf{b}$ has non-zero values only at the end points; the contributions from intermediate nodal points vanish because $p(x)y'(x)$ is continuous within the interval (a, b). Furthermore, if $y'(a) = y'(b) = 0$ from the boundary conditions, then the $\mathbf{b}$ vector becomes the zero vector. On the other hand, if $y(a) = y(b) = 0$, then $y_1 = y_5 = 0$ and the eigenvalue problem involves a 3×3 matrix with the unknowns y_2, y_3, and y_4. Similarly, if $y'(a) = y(b) = 0$, then we have a 4×4 matrix with the unknowns y_1, y_2, y_3, y_4, and $y_5 = 0$. Finally, if

Table 6.6.2: The Lowest Eigenvalue for Equation 6.6.30, Which Is Solved Using a Finite Element Method

L	λ
0.250	10.6745
0.100	10.2335
0.050	10.1717
0.020	10.1544
0.010	10.1520
0.002	10.1512

$y(a) = y'(b) = 0$, we again have a 4×4 matrix involving $y_1 = 0$ and the unknowns y_2, y_3, y_4, and y_5.

To illustrate this scheme, consider the Sturm-Liouville problem

$$y'' + (\lambda - x^2)y = 0, \qquad 0 < x < 1 \tag{6.6.30}$$

with $y(0) = y(1) = 0$. Here $p(x) = 1$, $q(x) = x^2$, and $r(x) = 1$.

The MATLAB begins with the choice of the number of elements, `N`. Once that is done, L immediately follows because `L = 1/(N-1)`. We will also need to have the value of x at the node points `x(n) = L*(n-1)` where n = 1:N.

With these preliminaries out of the way, we begin by setting up the matrices K and M given by Equation 6.6.27 and Equation 6.6.28. The corresponding MATLAB code is

```
for i = 1:N-1
  x_c = 0.5*(x(i) + x(i+1));
  p = 1; q = x_c*x_c; r = 1;
  K_11 = p/L + q*L/3; K_22 = K_11;
  K_12 = -p/L + q*L/6; K_21 = K_12;
  M_11 = r*L/3; M_22 = M_11;
  M_12 = r*L/6; M_21 = M_12;
  KK( i , i ) = KK( i , i ) + K_11;
  KK( i ,i+1) = KK( i ,i+1) + K_12;
  KK(i+1, i ) = KK(i+1, i ) + K_21;
  KK(i+1,i+1) = KK(i+1,i+1) + K_22;
  MM( i , i ) = MM( i , i ) + M_11;
  MM( i ,i+1) = MM( i ,i+1) + M_12;
  MM(i+1, i ) = MM(i+1, i ) + M_21;
  MM(i+1,i+1) = MM(i+1,i+1) + M_22;
end
```

Note that the arrays `KK` and `MM` have already been defined as $N \times N$ arrays with all of their elements set to zero.

Finally, because y_1 and y_N are zero, we must extract that portion of K and M for which $y_m \neq 0$. This is done as follows:

```
for j = 1:N-2
for i = 1:N-2
  A(i,j) = KK(i+1,j+1);
  B(i,j) = MM(i+1,j+1);
end; end
```

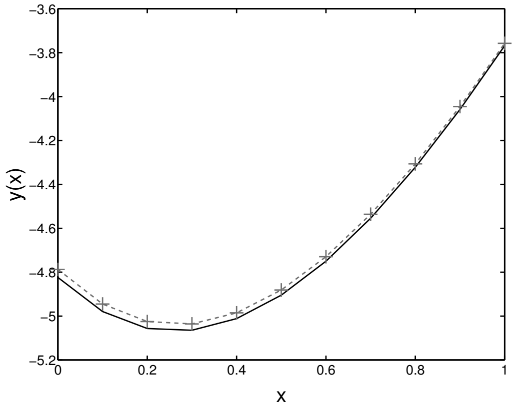


Figure 6.6.1: Numerical solution of $y'' + 2y' + y = 2x + 3\sin(x)$ with $y'(0) = -2$ and $y'(1) = 3$ using finite elements with $\Delta x = 0.1$. The crosses indicate the exact solution.

Finally, the eigenvalues are found by `eig(A,B)`. If the corresponding eigenfunction is desired, then the corresponding eigenfunction gives y_j for $j = 2, 3, \ldots, N-1$ using Equation 6.6.4. Table 6.6.2 illustrates how the lowest eigenvalue for Equation 6.6.30 improves in accuracy as the number of nodes is increased.

Project: Finite Element Solution of Boundary-Value Problems

In addition to solving the Sturm-Liouville problem, finite element methods can be used to solve the standard boundary-value problem:

$$-\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)y = f(x), \qquad 0 < x < 1,$$

where we specify $y(0)$ *or* $y'(0)$ at $x = 0$ and $y(1)$ *or* $y'(1)$ at $x = 1$. Although you could create your MATLAB code to solve this problem, MATLAB code has already been developed and is available online.[31] The purpose of this project is for you to become comfortable using this scheme.

Step 1: Using the method of undetermined coefficients, solve the boundary-value problem

$$y'' + 2y' + y = 2x + 3\sin(x), \qquad y'(0) = -2, \quad y'(1) = 3.$$

Show that

$$y(x) = 2x - 4 - \tfrac{3}{2}\cos(x) + \tfrac{1}{2}\left[3\sin(1) - 2\right]e^{1-x} + \tfrac{1}{2}\left[3\sin(1) - 2 - 8/e\right]xe^{1-x}.$$

Step 2: Show that the ordinary differential equation can be written as

$$\frac{d}{dx}\left(e^{2x}\frac{dy}{dx}\right) + e^{2x}y = 2xe^{2x} + 3e^{2x}\sin(x), \qquad y'(0) = -2, \quad y'(1) = 3.$$

Step 3: Find the numerical solution of the boundary-value problem using finite elements. Figure 6.6.1 illustrates the solution.

[31] For example, `http://people.sc.fsu.edu/~burkardt/m_src/fem1d/fem1d.html`

Further Readings

Hobson, E. W., 1965: *The Theory of Spherical and Ellipsoidal Harmonics.* Chelsea Publishers, 500 pp. The classic treatise on Legendre polynomials.

Lebedev, N. N., 1972: *Special Functions and Their Applications*. Dover, 308 pp. A very practical guide to the special functions found in the natural sciences and engineering.

Titchmarsh, E. C., 1946: *Eigenfunction Expansions Associated with Second Order Differential Equations.* Camp Press, 188 pp. A rigorous treatment of the Sturm-Liouville problem.

Watson, G. N., 1966: *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, 804 pp. The standard reference on Bessel functions.

Chapter 7
The Wave Equation

In this chapter we will study problems associated with the equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \tag{7.0.1}$$

where $u = u(x,t)$, x and t are the two independent variables, and c is a constant. This equation, called the *wave equation*, serves as the prototype for a wider class of *hyperbolic equations*

$$a(x,t)\frac{\partial^2 u}{\partial x^2} + b(x,t)\frac{\partial^2 u}{\partial x \partial t} + c(x,t)\frac{\partial^2 u}{\partial t^2} = f\left(x, t, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial t}\right), \tag{7.0.2}$$

where $b^2 > 4ac$. It arises in the study of many important physical problems involving wave propagation, such as the transverse vibrations of an elastic string and the longitudinal vibrations or torsional oscillations of a rod.

7.1 THE VIBRATING STRING

The motion of a string of length L and constant density ρ (mass per unit *length*) is a simple example of a physical system described by the wave equation. See Figure 7.1.1. Assuming that the equilibrium position of the string and the interval $[0, L]$ along the x-axis

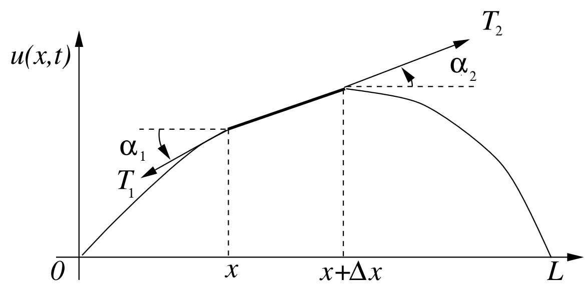


Figure 7.1.1: The vibrating string.

coincide, the equation of motion that describes the vertical displacement $u(x,t)$ of the string follows by considering a short piece whose ends are at x and $x+\Delta x$ and applying Newton's second law.

If we assume that the string is perfectly flexible and offers no resistance to bending, Figure 7.1.1 shows the forces on an element of the string. Applying Newton's second law in the x-direction, the sum of forces equals

$$-T(x)\cos(\alpha_1)+T(x+\Delta x)\cos(\alpha_2), \tag{7.1.1}$$

where $T(x)$ denotes the tensile force. If we assume that a point on the string moves only in the vertical direction, the sum of forces in Equation 7.1.1 equals zero and the horizontal component of tension is constant:

$$-T(x)\cos(\alpha_1)+T(x+\Delta x)\cos(\alpha_2)=0, \tag{7.1.2}$$

and

$$T(x)\cos(\alpha_1)=T(x+\Delta x)\cos(\alpha_2)=T, \text{ a constant.} \tag{7.1.3}$$

If gravity is the only external force, Newton's law in the vertical direction gives

$$-T(x)\sin(\alpha_1)+T(x+\Delta x)\sin(\alpha_2)-mg=m\frac{\partial^2 u}{\partial t^2}, \tag{7.1.4}$$

where u_{tt} is the acceleration. Because

$$T(x)=\frac{T}{\cos(\alpha_1)}, \quad \text{and} \quad T(x+\Delta x)=\frac{T}{\cos(\alpha_2)}, \tag{7.1.5}$$

then

$$-T\tan(\alpha_1)+T\tan(\alpha_2)-\rho g\Delta x=\rho\Delta x\frac{\partial^2 u}{\partial t^2}. \tag{7.1.6}$$

The quantities $\tan(\alpha_1)$ and $\tan(\alpha_2)$ equal the slope of the string at x and $x+\Delta x$, respectively; that is,

$$\tan(\alpha_1)=\frac{\partial u(x,t)}{\partial x}, \quad \text{and} \quad \tan(\alpha_2)=\frac{\partial u(x+\Delta x,t)}{\partial x}. \tag{7.1.7}$$

Substituting Equation 7.1.7 into Equation 7.1.6,

$$T\left[\frac{\partial u(x+\Delta x,t)}{\partial x}-\frac{\partial u(x,t)}{\partial x}\right]=\rho\Delta x\left(\frac{\partial^2 u}{\partial t^2}+g\right). \tag{7.1.8}$$

After dividing through by Δx, we have a difference quotient on the left:

$$\frac{T}{\Delta x}\left[\frac{\partial u(x+\Delta x,t)}{\partial x} - \frac{\partial u(x,t)}{\partial x}\right] = \rho\left(\frac{\partial^2 u}{\partial t^2} + g\right). \tag{7.1.9}$$

In the limit as $\Delta x \to 0$, this difference quotient becomes a partial derivative with respect to x, leaving Newton's second law in the form

$$T\frac{\partial^2 u}{\partial x^2} = \rho\frac{\partial^2 u}{\partial t^2} + \rho g, \tag{7.1.10}$$

or

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} + \frac{g}{c^2}, \tag{7.1.11}$$

where $c^2 = T/\rho$. Because u_{tt} is generally much larger than g, we can neglect the last term, giving the equation of the vibrating string as

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}. \tag{7.1.12}$$

Equation 7.1.12 is the one-dimensional *wave equation.*

As a second example[1] we derive the threadline equation, which describes how a thread composed of yarn vibrates as we draw it between two eyelets spaced a distance L apart. We assume that the tension in the thread is constant, the vibrations are small, the thread is perfectly flexible, the effects of gravity and air drag are negligible, and the mass of the thread per unit length is constant. Unlike the vibrating string between two fixed ends, we draw the threadline through the eyelets at a speed V so that a segment of thread experiences motion in both the x and y directions as it vibrates about its equilibrium position. The eyelets may move in the vertical direction.

From Newton's second law,

$$\frac{d}{dt}\left(m\frac{dy}{dt}\right) = \sum \text{ forces}, \tag{7.1.13}$$

where m is the mass of the thread. But

$$\frac{dy}{dt} = \frac{\partial y}{\partial t} + \frac{dx}{dt}\frac{\partial y}{\partial x}. \tag{7.1.14}$$

Because $dx/dt = V$,

$$\frac{dy}{dt} = \frac{\partial y}{\partial t} + V\frac{\partial y}{\partial x}, \tag{7.1.15}$$

and

$$\frac{d}{dt}\left(m\frac{dy}{dt}\right) = \frac{\partial}{\partial t}\left[m\left(\frac{\partial y}{\partial t} + V\frac{\partial y}{\partial x}\right)\right] + V\frac{\partial}{\partial x}\left[m\left(\frac{\partial y}{\partial t} + V\frac{\partial y}{\partial x}\right)\right]. \tag{7.1.16}$$

Because both m and V are constant, it follows that

$$\frac{d}{dt}\left(m\frac{dy}{dt}\right) = m\frac{\partial^2 y}{\partial t^2} + 2mV\frac{\partial^2 y}{\partial x \partial t} + mV^2\frac{\partial^2 y}{\partial x^2}. \tag{7.1.17}$$

[1] See Swope, R. D., and W. F. Ames, 1963: Vibrations of a moving threadline. *J. Franklin Inst.*, **275**, 36–55.

The sum of the forces again equals

$$T\frac{\partial^2 y}{\partial x^2}\Delta x \tag{7.1.18}$$

so that the threadline equation is

$$T\frac{\partial^2 y}{\partial x^2}\Delta x = m\frac{\partial^2 y}{\partial t^2} + 2mV\frac{\partial^2 y}{\partial x \partial t} + mV^2\frac{\partial^2 y}{\partial x^2}, \tag{7.1.19}$$

or

$$\frac{\partial^2 y}{\partial t^2} + 2V\frac{\partial^2 y}{\partial x \partial t} + \left(V^2 - \frac{gT}{\rho}\right)\frac{\partial^2 y}{\partial x^2} = 0, \tag{7.1.20}$$

where ρ is the density of the thread. Although Equation 7.1.20 is *not* the classic wave equation given in Equation 7.1.12, it is an example of a hyperbolic equation. As we shall see, the solutions to hyperbolic equations share the same behavior, namely, wave-like motion.

7.2 INITIAL CONDITIONS: CAUCHY PROBLEM

Any mathematical model of a physical process must include not only the governing differential equation but also any conditions that are imposed on the solution. For example, in time-dependent problems the solution must conform to the initial condition of the modeled process. Finding those solutions that satisfy the initial conditions (initial data) is called the *Cauchy problem*.

In the case of partial differential equations with second-order derivatives in time, such as the wave equation, we correctly pose the *Cauchy boundary condition* if we specify the value of the solution $u(x, t_0) = f(t)$ *and* its time derivative $u_t(x, t_0) = g(t)$ at some initial time t_0, usually taken to be $t_0 = 0$. The functions $f(t)$ and $g(t)$ are called the *Cauchy data*. We require two conditions involving time because the differential equation has two time derivatives.

In addition to the initial conditions, we must specify boundary conditions in the spatial direction. For example, we may require that the end of the string be fixed. In the next chapter, we discuss boundary conditions in greater depth. However, one boundary condition that is uniquely associated with the wave equation on an open domain is the *radiation condition*. It requires that the waves radiate off to infinity and remain finite as they propagate there.

In summary, Cauchy boundary conditions, along with the appropriate spatial boundary conditions, uniquely determine the solution to the wave equation; any additional information is extraneous. Having developed the differential equation and initial conditions necessary to solve the wave equation, let us now turn to the actual methods used to solve this equation.

7.3 SEPARATION OF VARIABLES

Separation of variables is the most popular method for solving the wave equation. Despite its current widespread use, its initial application to the vibrating string problem was controversial because of the use of a half-range Fourier sine series to represent the initial conditions. On one side, Daniel Bernoulli claimed (in 1775) that he could represent any general initial condition with this technique. To d'Alembert and Euler, however, the half-range Fourier sine series, with its period of $2L$, could not possibly represent any arbitrary function.[2] However, by 1807 Bernoulli was proven correct by the use of separation of variables in the heat conduction problem and it rapidly grew in acceptance.[3] In the following examples we show how to apply this method.

Separation of variables consists of four distinct steps that convert a second-order partial differential equation into two ordinary differential equations. First, we *assume* that the solution equals the product $X(x)T(t)$. Direct substitution into the partial differential equation and boundary conditions yields two ordinary differential equations and the corresponding boundary conditions. Step two involves solving a boundary-value problem of the Sturm-Liouville type. In step three we find the corresponding time dependence. Finally we construct the complete solution as a sum of all product solutions. Upon applying the initial conditions, we have an eigenfunction expansion and must compute the Fourier coefficients. The substitution of these coefficients into the summation yields the complete solution.

• Example 7.3.1

Let us solve the wave equation for the special case when we clamp the string at $x = 0$ and $x = L$. Mathematically, we find the solution to the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < L, \quad 0 < t, \tag{7.3.1}$$

which satisfies the initial conditions

$$u(x, 0) = f(x), \quad \frac{\partial u(x, 0)}{\partial t} = g(x), \quad 0 < x < L, \tag{7.3.2}$$

and the boundary conditions

$$u(0, t) = u(L, t) = 0, \quad 0 < t. \tag{7.3.3}$$

For the present, we leave the Cauchy data quite arbitrary.

We begin by assuming that the solution $u(x, t)$ equals the product $X(x)T(t)$. (Here T no longer denotes tension.) Because

$$\frac{\partial^2 u}{\partial t^2} = X(x)T''(t), \tag{7.3.4}$$

and

$$\frac{\partial^2 u}{\partial x^2} = X''(x)T(t), \tag{7.3.5}$$

the wave equation becomes

$$c^2 X''T = T''X, \tag{7.3.6}$$

or

$$\frac{X''}{X} = \frac{T''}{c^2 T}, \tag{7.3.7}$$

---

[2] See Hobson, E. W., 1957: *The Theory of Functions of a Real Variable and the Theory of Fourier's Series, Vol. 2*. Dover Publishers, Sections 312–314.

[3] Lützen, J., 1984: Sturm and Liouville's work on ordinary linear differential equations. The emergence of Sturm-Liouville theory. *Arch. Hist. Exact Sci.*, **29**, 317.

after dividing through by $c^2X(x)T(t)$. Because the left side of Equation 7.3.7 depends only on x and the right side depends only on t, both sides must equal a constant. We write this separation constant $-\lambda$ and separate Equation 7.3.7 into two ordinary differential equations:

$$T'' + c^2\lambda T = 0, \qquad 0 < t, \tag{7.3.8}$$

and

$$X'' + \lambda X = 0, \qquad 0 < x < L. \tag{7.3.9}$$

We now rewrite the boundary conditions in terms of $X(x)$ by noting that the boundary conditions become

$$u(0,t) = X(0)T(t) = 0, \tag{7.3.10}$$

and

$$u(L,t) = X(L)T(t) = 0 \tag{7.3.11}$$

for $0 < t$. If we were to choose $T(t) = 0$, then we would have a trivial solution for $u(x,t)$. Consequently,

$$X(0) = X(L) = 0. \tag{7.3.12}$$

This concludes the first step.

In the second step we consider three possible values for λ: $\lambda < 0$, $\lambda = 0$, and $\lambda > 0$. Turning first to $\lambda < 0$, we set $\lambda = -m^2$ so that square roots of λ will not appear later on and m is real. The general solution of Equation 7.3.9 is

$$X(x) = A\cosh(mx) + B\sinh(mx). \tag{7.3.13}$$

Because $X(0) = 0$, $A = 0$. On the other hand, $X(L) = B\sinh(mL) = 0$. The function $\sinh(mL)$ does not equal to zero since $mL \neq 0$ (recall $m > 0$). Thus, $B = 0$ and we have trivial solutions for a positive separation constant.

If $\lambda = 0$, the general solution now becomes

$$X(x) = C + Dx. \tag{7.3.14}$$

The condition $X(0) = 0$ yields $C = 0$ while $X(L) = 0$ yields $DL = 0$ or $D = 0$. Hence, we have a trivial solution for the $\lambda = 0$ separation constant.

If $\lambda = k^2 > 0$, the general solution to Equation 7.3.9 is

$$X(x) = E\cos(kx) + F\sin(kx). \tag{7.3.15}$$

The condition $X(0) = 0$ results in $E = 0$. On the other hand, $X(L) = F\sin(kL) = 0$. If we wish to avoid a trivial solution in this case ($F \neq 0$), $\sin(kL) = 0$, or $k_n = n\pi/L$, and $\lambda_n = n^2\pi^2/L^2$. The x-dependence equals $X_n(x) = F_n\sin(n\pi x/L)$. We added the n subscript to k and λ to indicate that these quantities depend on n. This concludes the second step.

Turning to Equation 7.3.8 for the third step, the solution to the $T(t)$ equation is

$$T_n(t) = G_n\cos(k_nct) + H_n\sin(k_nct), \tag{7.3.16}$$

where G_n and H_n are arbitrary constants. For each $n = 1, 2, 3, \ldots$, a particular solution that satisfies the wave equation and prescribed boundary conditions is

$$u_n(x,t) = F_n\sin\left(\frac{n\pi x}{L}\right)\left[G_n\cos\left(\frac{n\pi ct}{L}\right) + H_n\sin\left(\frac{n\pi ct}{L}\right)\right], \tag{7.3.17}$$

or

$$u_n(x,t) = \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right)\right], \qquad \textbf{(7.3.18)}$$

where $A_n = F_nG_n$ and $B_n = F_nH_n$. This concludes the third step.

An equivalent method of finding the product solution is to treat Equation 7.3.9 along with $X(0) = X(L) = 0$ as a Sturm-Liouville problem. In this method we obtain the spatial dependence by solving the Sturm-Liouville problem and finding the corresponding eigenvalues λ_n and eigenfunctions. Next we solve for $T_n(t)$. Finally we form the product solution $u_n(x,t)$ by multiplying the eigenfunction times the temporal dependence.

For any choice of A_n and B_n, Equation 7.3.18 is a solution of the partial differential equation, Equation 7.3.1, also satisfying the boundary conditions, Equation 7.3.3. Therefore, any linear combination of $u_n(x,t)$ also satisfies the partial differential equation and the boundary conditions. In making this linear combination we need no new constants because A_n and B_n are still arbitrary. We have, then,

$$u(x,t) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(\frac{n\pi ct}{L}\right) + B_n \sin\left(\frac{n\pi ct}{L}\right)\right]. \qquad \textbf{(7.3.19)}$$

Our method of using particular solutions to build up the general solution illustrates the powerful *principle of linear superposition*, which is applicable to any *linear* system. This principle states that if u_1 and u_2 are any solutions of a linear homogeneous partial differential equation in any region, then $u = c_1u_1 + c_2u_2$ is also a solution of that equation in that region, where c_1 and c_2 are any constants. We can generalize this to an infinite sum. It is extremely important because it allows us to construct general solutions to partial differential equations from particular solutions to the same problem.

Our fourth and final task remains to determine A_n and B_n. At $t = 0$,

$$u(x,0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right) = f(x), \qquad \textbf{(7.3.20)}$$

and

$$u_t(x,0) = \sum_{n=1}^{\infty} \frac{n\pi c}{L} B_n \sin\left(\frac{n\pi x}{L}\right) = g(x). \qquad \textbf{(7.3.21)}$$

Both of these series are Fourier half-range sine expansions over the interval $(0, L)$. Applying the results from Section 5.3,

$$A_n = \frac{2}{L}\int_0^L f(x) \sin\left(\frac{n\pi x}{L}\right)\, dx, \qquad \textbf{(7.3.22)}$$

and

$$\frac{n\pi c}{L} B_n = \frac{2}{L}\int_0^L g(x) \sin\left(\frac{n\pi x}{L}\right)\, dx, \qquad \textbf{(7.3.23)}$$

or

$$B_n = \frac{2}{n\pi c}\int_0^L g(x) \sin\left(\frac{n\pi x}{L}\right)\, dx. \qquad \textbf{(7.3.24)}$$

At this point we might ask ourselves whether the Fourier series solution to the wave equation always converges. For the case $g(x) = 0$, Carslaw[4] showed that if the initial position of the

[4] Carslaw, H. S., 1902: Note on the use of Fourier's series in the problem of the transverse vibrations of strings. *Proc. Edinburgh Math. Soc., Ser. 1*, **20**, 23–28.

string forms a curve so that $f(x)$ or the slope $f'(x)$ is continuous between $x = 0$ and $x = L$, then the series converges uniformly.

As an example, let us take the initial conditions

$$f(x) = \begin{cases} 0, & 0 < x \le L/4, \\ 4h\left(\frac{x}{L} - \frac{1}{4}\right), & L/4 \le x \le L/2, \\ 4h\left(\frac{3}{4} - \frac{x}{L}\right), & L/2 \le x \le 3L/4, \\ 0, & 3L/4 \le x < L, \end{cases} \tag{7.3.25}$$

and

$$g(x) = 0, \qquad 0 < x < L. \tag{7.3.26}$$

In this particular example, $B_n = 0$ for all n because $g(x) = 0$. On the other hand,

$$A_n = \frac{8h}{L}\int_{L/4}^{L/2}\left(\frac{x}{L} - \frac{1}{4}\right)\sin\left(\frac{n\pi x}{L}\right)dx + \frac{8h}{L}\int_{L/2}^{3L/4}\left(\frac{3}{4} - \frac{x}{L}\right)\sin\left(\frac{n\pi x}{L}\right)dx \tag{7.3.27}$$

$$= \frac{8h}{n^2\pi^2}\left[2\sin\left(\frac{n\pi}{2}\right) - \sin\left(\frac{3n\pi}{4}\right) - \sin\left(\frac{n\pi}{4}\right)\right] \tag{7.3.28}$$

$$= \frac{8h}{n^2\pi^2}\left[2\sin\left(\frac{n\pi}{2}\right) - 2\sin\left(\frac{n\pi}{2}\right)\cos\left(\frac{n\pi}{4}\right)\right] \tag{7.3.29}$$

$$= \frac{16h}{n^2\pi^2}\sin\left(\frac{n\pi}{2}\right)\left[1 - \cos\left(\frac{n\pi}{4}\right)\right] = \frac{32h}{n^2\pi^2}\sin\left(\frac{n\pi}{2}\right)\sin^2\left(\frac{n\pi}{8}\right), \tag{7.3.30}$$

because $\sin(A) + \sin(B) = 2\sin[\frac{1}{2}(A+B)]\cos[\frac{1}{2}(A-B)]$, and $1 - \cos(2A) = 2\sin^2(A)$. Therefore,

$$u(x,t) = \frac{32h}{\pi^2}\sum_{n=1}^{\infty}\sin\left(\frac{n\pi}{2}\right)\sin^2\left(\frac{n\pi}{8}\right)\frac{1}{n^2}\sin\left(\frac{n\pi x}{L}\right)\cos\left(\frac{n\pi ct}{L}\right). \tag{7.3.31}$$

Because $\sin(n\pi/2)$ vanishes for n even, so does A_n. If Equation 7.3.31 were evaluated on a computer, considerable time and effort would be wasted. Consequently it is preferable to rewrite Equation 7.3.31 so that we eliminate these vanishing terms. The most convenient method introduces the general expression $n = 2m - 1$ for any odd integer, where $m = 1, 2, 3, \ldots$, and notes that $\sin[(2m-1)\pi/2] = (-1)^{m+1}$. Therefore, Equation 7.3.31 becomes

$$u(x,t) = \frac{32h}{\pi^2}\sum_{m=1}^{\infty}\frac{(-1)^{m+1}}{(2m-1)^2}\sin^2\left[\frac{(2m-1)\pi}{8}\right]\sin\left[\frac{(2m-1)\pi x}{L}\right]\cos\left[\frac{(2m-1)\pi ct}{L}\right]. \tag{7.3.32}$$

Although we completely solved the problem, it is useful to rewrite Equation 7.3.32 as

$$u(x,t) = \frac{1}{2}\sum_{n=1}^{\infty}A_n\left\{\sin\left[\frac{n\pi}{L}(x - ct)\right] + \sin\left[\frac{n\pi}{L}(x + ct)\right]\right\} \tag{7.3.33}$$

through the application of the trigonometric identity $\sin(A)\cos(B) = \frac{1}{2}\sin(A - B) + \frac{1}{2}\sin(A + B)$. From general physics we find expressions like $\sin[k_n(x - ct)]$ or $\sin(kx - \omega t)$ arising in studies of simple wave motions. The quantity $\sin(kx - \omega t)$ is the mathematical description of a propagating wave in the sense that we must move to the right at the speed c if we wish to keep in the same position relative to the nearest crest and trough. The

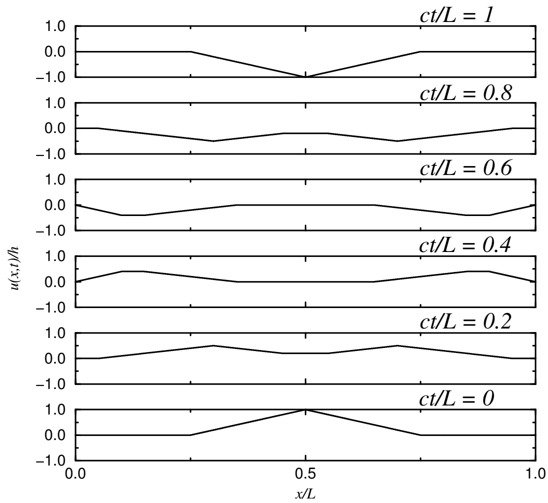


Figure 7.3.1: The vibration of a string $u(x,t)/h$ at various positions x/L at the times $ct/L = 0$, 0.2, 0.4, 0.6, 0.8, and 1. For times $1 < ct/L < 2$ the pictures appear in reverse time order.

quantities k, ω, and c are the wavenumber, frequency, and phase speed or wave velocity, respectively. The relationship $\omega = kc$ holds between the frequency and phase speed.

It may seem paradoxical that we are talking about traveling waves in a problem dealing with waves confined on a string of length L. Actually we are dealing with standing waves because at the same time that a wave is propagating to the right its mirror image is running to the left so that there is no resultant progressive wave motion. Figures 7.3.1 and 7.3.2 illustrate our solution. Figure 7.3.1 gives various cross sections. The single large peak at $t = 0$ breaks into two smaller peaks that race towards the two ends. At each end, they reflect and turn upside down as they propagate back towards $x = L/2$ at $ct/L = 1$. This large, negative peak at $x = L/2$ again breaks apart, with the two smaller peaks propagating towards the endpoints. They reflect and again become positive peaks as they propagate back to $x = L/2$ at $ct/L = 2$. After that time, the whole process repeats itself.

MATLAB can be used to examine the solution in its totality. The script

```
% set parameters for the calculation
clear; M = 50; dx = 0.02; dt = 0.02;
% compute Fourier coefficients
sign = 32;
for m = 1:M
  temp1 = (2*m-1)*pi; temp2 = sin(temp1/8);
  a(m) = sign * temp2 * temp2 / (temp1 * temp1);
  sign = -sign;
end
% compute grid and initialize solution
X = [0:dx:1]; T = [0:dt:2];
u = zeros(length(T),length(X));
XX = repmat(X,[length(T) 1]);
TT = repmat(T',[1 length(X)]);
% compute solution from Equation 7.3.32
for m = 1:M
  temp1 = (2*m-1)*pi;
  u = u + a(m) .* sin(temp1*XX) .* cos(temp1*TT);
```

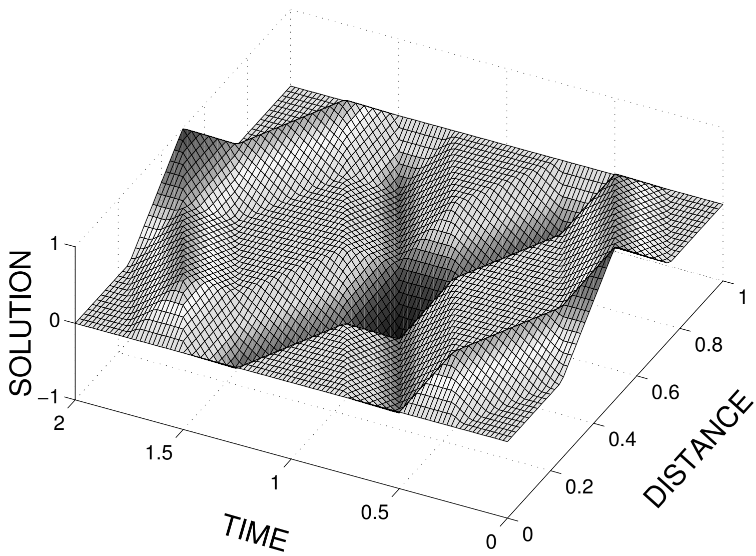


Figure 7.3.2: Two-dimensional plot of the vibration of a string $u(x,t)/h$ at various times ct/L and positions x/L.

```
end
% plot space/time picture of the solution
surf(XX,TT,u)
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

gives a three-dimensional view of Equation 7.3.32. The solution can be viewed in many different prospects using the interactive capacity of MATLAB.

An important dimension to the vibrating string problem is the fact that the wavenumber k_n is not a free parameter but has been restricted to the values of $n\pi/L$. This restriction on wavenumber is common in wave problems dealing with limited domains (for example, a building, ship, lake, or planet) and these oscillations are given the special name of *normal modes* or *natural vibrations*.

In our problem of the vibrating string, all of the components propagate with the same phase speed. That is, all of the waves, regardless of wavenumber k_n, move the characteristic distance $c\Delta t$ or $-c\Delta t$ after the time interval Δt elapsed. In the next example we will see that this is not always true. □

• **Example 7.3.2: Dispersion**

In the preceding example, the solution to the vibrating string problem consisted of two simple waves, each propagating with a phase speed c to the right and left. In many problems where the equations of motion are a little more complicated than Equation 7.3.1, all of the harmonics no longer propagate with the same phase speed but at a speed that depends upon the wavenumber. In such systems the phase relation varies between the harmonics and these systems are referred to as *dispersive*.

A modification of the vibrating string problem provides a simple illustration. We now subject each element of the string to an additional applied force that is proportional to its displacement:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} - hu, \quad 0 < x < L, \quad 0 < t, \tag{7.3.34}$$

where $h > 0$ is constant. For example, if we embed the string in a thin sheet of rubber, then in addition to the restoring force due to tension, there is a restoring force due to the rubber on each portion of the string. From its use in the quantum mechanics of "scalar" mesons, Equation 7.3.34 is often referred to as the *Klein-Gordon* equation.

We shall again look for particular solutions of the form $u(x,t) = X(x)T(t)$. This time, however,

$$XT'' - c^2X''T + hXT = 0, \tag{7.3.35}$$

or

$$\frac{T''}{c^2T} + \frac{h}{c^2} = \frac{X''}{X} = -\lambda, \tag{7.3.36}$$

which leads to two ordinary differential equations

$$X'' + \lambda X = 0, \tag{7.3.37}$$

and

$$T'' + (\lambda c^2 + h)T = 0. \tag{7.3.38}$$

If we attach the string at $x = 0$ and $x = L$, the $X(x)$ solution is

$$X_n(x) = \sin\left(\frac{n\pi x}{L}\right) \tag{7.3.39}$$

with $k_n = n\pi/L$, and $\lambda_n = n^2\pi^2/L^2$. On the other hand, the $T(t)$ solution becomes

$$T_n(t) = A_n \cos\left(\sqrt{k_n^2c^2 + h}\, t\right) + B_n \sin\left(\sqrt{k_n^2c^2 + h}\, t\right), \tag{7.3.40}$$

so that the product solution is

$$u_n(x,t) = \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(\sqrt{k_n^2c^2 + h}\, t\right) + B_n \sin\left(\sqrt{k_n^2c^2 + h}\, t\right)\right]. \tag{7.3.41}$$

Finally, the general solution becomes

$$u(x,t) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(\sqrt{k_n^2c^2 + h}\, t\right) + B_n \sin\left(\sqrt{k_n^2c^2 + h}\, t\right)\right] \tag{7.3.42}$$

from the principle of linear superposition. Let us consider the case when $B_n = 0$. Then we can write Equation 7.3.42 as

$$u(x,t) = \sum_{n=1}^{\infty} \frac{A_n}{2}\left[\sin\left(k_nx + \sqrt{k_n^2c^2 + h}\, t\right) + \sin\left(k_nx - \sqrt{k_n^2c^2 + h}\, t\right)\right]. \tag{7.3.43}$$

Comparing our results with Equation 7.3.33, the distance that a particular mode k_n moves during the time interval Δt depends not only upon external parameters such as h, the tension and density of the string, but also upon its wavenumber (or equivalently, wavelength). Furthermore, the frequency of a particular harmonic is larger than that when $h = 0$. This result is not surprising, because the added stiffness of the medium should increase the natural frequencies.

The importance of dispersion lies in the fact that if the solution $u(x,t)$ is a superposition of progressive waves in the same direction, then the phase relationship between the different harmonics changes with time. Because most signals consist of an infinite series

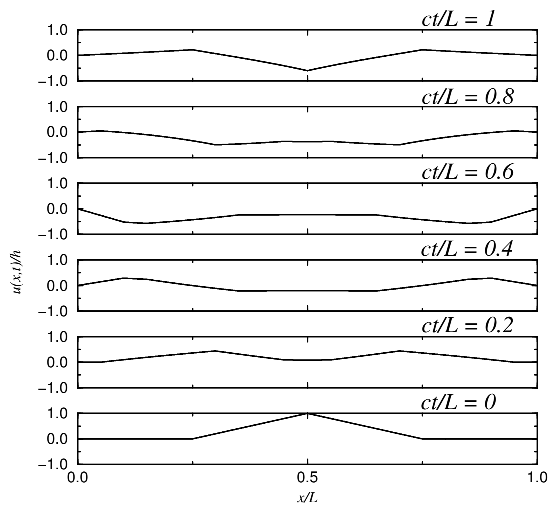


Figure 7.3.3: The vibration of a string $u(x,t)/h$ embedded in a thin sheet of rubber at various positions x/L at the times $ct/L = 0$, 0.2, 0.4, 0.6, 0.8, and 1 for $hL^2/c^2 = 10$. The same parameters were used as in Figure 7.3.1.

of these progressive waves, dispersion causes the signal to become garbled. We show this by comparing the solution, Equation 7.3.42 given in Figures 7.3.3 and 7.3.4 for the initial conditions, Equation 7.3.25 and Equation 7.3.26, with $hL^2/c^2 = 10$, to the results given in Figures 7.3.1 and 7.3.2. In the case of Figure 7.3.4, the MATLAB script line

```
u = u + a(m) .* sin(temp1*XX) .* cos(temp1*TT);
```

has been replaced with

```
temp2 = temp1 * sqrt(1 + H/(temp1*temp1));
u = u + a(m) .* sin(temp1*XX) .* cos(temp2*TT);
```

where `H = 10` is defined earlier in the script. Note how garbled the picture becomes at $ct/L = 2$ in Figure 7.3.4 compared to the nondispersive solution at the same time in Figure 7.3.2. □

• Example 7.3.3: Damped wave equation

In the previous example a slight modification of the wave equation resulted in a wave solution where each Fourier harmonic propagates with its own particular phase speed. In this example we introduce a modification of the wave equation that results not only in dispersive waves but also in the exponential decay of the amplitude as the wave propagates.

So far we neglected the reaction of the surrounding medium (air or water, for example) on the motion of the string. For small-amplitude motions this reaction opposes the motion of each element of the string and is proportional to the element's velocity. The equation of motion, when we account for the tension and friction in the medium but not its stiffness or internal friction, is

$$\frac{\partial^2 u}{\partial t^2} + 2h\frac{\partial u}{\partial t} = c^2\frac{\partial^2 u}{\partial x^2}, \quad 0 < x < L, \quad 0 < t. \qquad \textbf{(7.3.44)}$$

Because Equation 7.3.44 first arose in the mathematical description of the telegraph,[5] it is

[5] The first published solution was by Kirchhoff, G., 1857: Über die Bewegung der Electrität in Drähten. *Ann. Phys. Chem.*, **100**, 193–217. English translation: Kirchhoff, G., 1857: On the motion of electricity in wires. *Philos. Mag., Ser. 4*, **13**, 393–412.

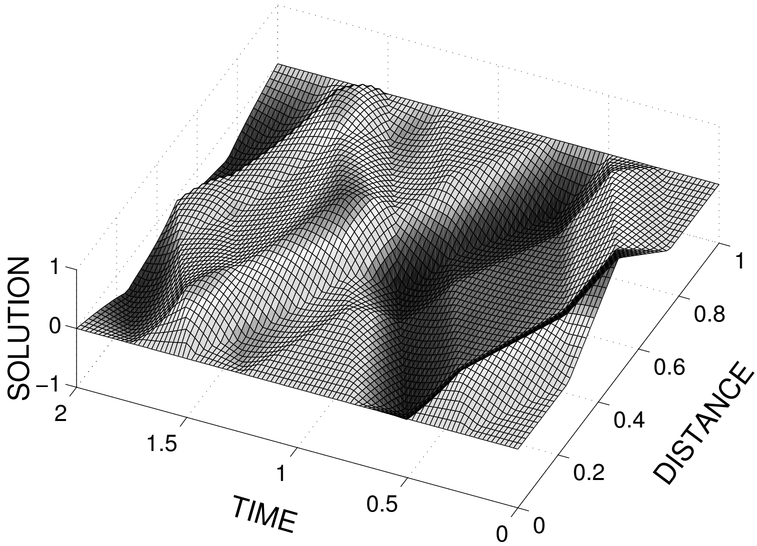


Figure 7.3.4: The two-dimensional plot of the vibration of a string $u(x,t)/h$ embedded in a thin sheet of rubber at various times ct/L and positions x/L for $hL^2/c^2 = 10$.

generally known as the *equation of telegraphy*. The effect of friction is, of course, to damp out the free vibration.

Let us assume a solution of the form $u(x,t) = X(x)T(t)$ and separate the variables to obtain the two ordinary differential equations:

$$X'' + \lambda X = 0, \tag{7.3.45}$$

and

$$T'' + 2hT' + \lambda c^2 T = 0 \tag{7.3.46}$$

with $X(0) = X(L) = 0$. Friction does not affect the shape of the normal modes; they are still

$$X_n(x) = \sin\left(\frac{n\pi x}{L}\right) \tag{7.3.47}$$

with $k_n = n\pi/L$ and $\lambda_n = n^2\pi^2/L^2$.

The solution for the $T(t)$ equation is

$$T_n(t) = e^{-ht}\left[A_n \cos\left(\sqrt{k_n^2c^2 - h^2}\, t\right) + B_n \sin\left(\sqrt{k_n^2c^2 - h^2}\, t\right)\right] \tag{7.3.48}$$

with the condition that $k_n c > h$. If we violate this condition, the solutions are two exponentially decaying functions in time. Because most physical problems usually fulfill this condition, we concentrate on this solution.

From the principle of linear superposition, the general solution is

$$u(x,t) = e^{-ht}\sum_{n=1}^{\infty}\sin\left(\frac{n\pi x}{L}\right)\left[A_n \cos\left(\sqrt{k_n^2c^2 - h^2}\, t\right) + B_n \sin\left(\sqrt{k_n^2c^2 - h^2}\, t\right)\right], \tag{7.3.49}$$

where $\pi c > hL$. From Equation 7.3.49 we see two important effects. First, the presence of friction slows all of the harmonics. Furthermore, friction dampens all of the harmonics.

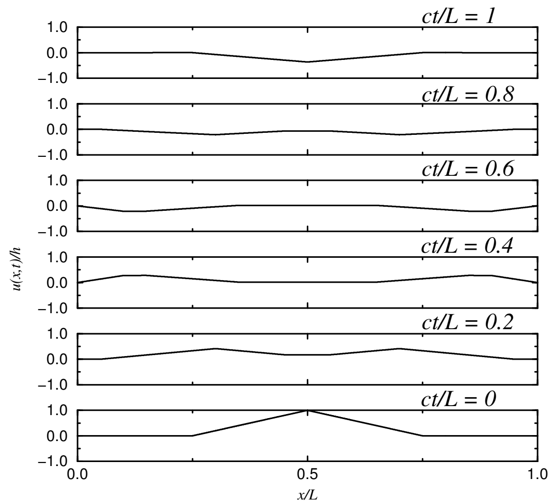


Figure 7.3.5: The vibration of a string $u(x,t)/h$ with frictional dissipation at various positions x/L at the times $ct/L = 0$, 0.2, 0.4, 0.6, 0.8, and 1 for $hL/c = 1$. The same parameters were used as in Figure 7.3.1.

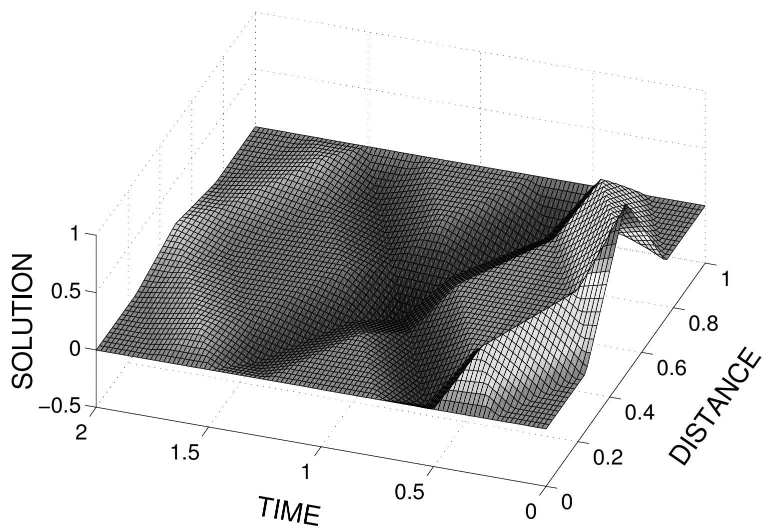


Figure 7.3.6: The vibration of a string $u(x,t)/h$ with frictional dissipation at various times ct/L and positions x/L for $hL/c = 1$.

Figures 7.3.5 and 7.3.6 illustrate the solution using the initial conditions given by Equation 7.3.25 and Equation 7.3.26 with $hL/c = 1$. In the case of Figure 7.3.6, the script line that produced Figure 7.3.2:

```
u = u + a(m) .* sin(temp1*XX) .* cos(temp1*TT);
```

has been replaced with

```
temp2 = temp1 * sqrt(1 - (H*H)/(temp1*temp1));
u = u + a(m) .* exp(-H*TT) .* sin(temp1*XX) .* cos(temp2*TT);
```

where `H = 1` is defined earlier in the script. Because this is a rather large coefficient of friction, Figures 7.3.5 and 7.3.6 exhibit rapid damping as well as dispersion.

This damping and dispersion of waves also occurs in solutions of the equation of telegraphy where the solutions are progressive waves. Because early telegraph lines were short, time delay effects were negligible. However, when engineers laid the first transoceanic cables

Table 7.3.1: Technological Innovation on Transatlantic Telegraph Cables

Year	Technological Innovation	Performance (words/min)
1857–58	Mirror galvanometer	3–7
1870	Condensers	12
1872	Siphon recorder	17
1879	Duplex	24
1894	Larger diameter cable	72–90
1915–20	Brown drum repeater and Heurtley magnifier	100
1923–28	Magnetically loaded lines	300–320
1928–32	Electronic signal shaping amplifiers and time division multiplexing	480
1950	Repeaters on the continental shelf	100–300
1956	Repeater telephone cables	21600

From Coates, V. T., and B. Finn, 1979: *A Retrospective Technology Assessment: Submarine Telegraphy. The Transatlantic Cable of 1866.* San Francisco Press, Inc., 268 pp.

in the 1850s, the time delay became seconds and differences in the velocity of propagation of different frequencies, as predicted by Equation 7.3.49, became noticeable to the operators. Table 7.3.1 gives the transmission rate for various transatlantic submarine telegraph lines. As it shows, increases in the transmission rates during the nineteenth century were due primarily to improvements in terminal technology.

When they instituted long-distance telephony just before the turn of the twentieth century, this difference in velocity between frequencies should have limited the circuits to a few tens of miles.[6] However, in 1899, Prof. Michael Pupin at Columbia University showed that by adding inductors ("loading coils") to the line at regular intervals the velocities at the different frequencies could be equalized.[7] Heaviside[8] and the French engineer Vaschy[9] made similar suggestions in the nineteenth century. Thus, adding resistance and inductance, which would seem to make things worse, actually made possible long-distance telephony. Today you can see these loading coils as you drive along the street; they are the black cylinders, approximately one between each pair of telephone poles, spliced into the telephone cable. The loading of long submarine telegraph cables had to wait for the development of permalloy and mu-metal materials of high magnetic induction. □

---

[6] Rayleigh, J. W., 1884: On telephoning through a cable. *Brit. Assoc. Rep.*, 632–633; Jordan, D. W., 1982: The adoption of self-induction by telephony, 1886–1889. *Ann. Sci.*, **39**, 433–461.

[7] There is considerable controversy concerning who is exactly the inventor. See Brittain, J. E., 1970: The introduction of the loading coil: George A. Campbell and Michael I. Pupin. *Tech. Culture*, **11**, 36–57.

[8] First published 3 June 1887. Reprinted in Heaviside, O., 1970: *Electrical Papers, Vol. 2*. Chelsea Publishing, pp. 119–124.

[9] See Devaux-Charbonnel, X. G. F., 1917: La contribution des ingénieurs français à la téléphonie à grande distance par câbles souterrains: Vaschy et Barbarat. *Rev. Gén. Électr.*, **2**, 288–295.

• **Example 7.3.4: Axisymmetric vibrations of a circular membrane**

The wave equation

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}, \quad 0 \le r < a, \quad 0 < t \tag{7.3.50}$$

governs axisymmetric vibrations of a circular membrane, where $u(r,t)$ is the vertical displacement of the membrane, r is the radial distance, t is time, c is the square root of the ratio of the tension of the membrane to its density, and a is the radius of the membrane. We will solve Equation 7.3.50 when the membrane is initially at rest, $u(r,0) = 0$, and struck so that its initial velocity is

$$\frac{\partial u(r,0)}{\partial t} = \begin{cases} P/(\pi\epsilon^2\rho), & 0 \le r < \epsilon, \\ 0, & \epsilon < r < a. \end{cases} \tag{7.3.51}$$

If this problem can be solved by separation of variables, then $u(r,t) = R(r)T(t)$. Following the substitution of this $u(r,t)$ into Equation 7.3.50, separation of variables leads to

$$\frac{1}{rR}\frac{d}{dr}\left(r\frac{dR}{dr}\right) = \frac{1}{c^2T}\frac{d^2T}{dt^2} = -k^2, \tag{7.3.52}$$

or

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dR}{dr}\right) + k^2R = 0, \tag{7.3.53}$$

and

$$\frac{d^2T}{dt^2} + k^2c^2T = 0. \tag{7.3.54}$$

The separation constant $-k^2$ must be negative so that we obtain solutions that remain bounded in the region $0 \le r < a$ and can satisfy the boundary condition. This boundary condition is $u(a,t) = R(a)T(t) = 0$, or $R(a) = 0$.

The solutions of Equation 7.3.53 and Equation 7.3.54, subject to the boundary condition, are

$$R_n(r) = J_0\left(\frac{\lambda_n r}{a}\right), \tag{7.3.55}$$

and

$$T_n(t) = A_n \sin\left(\frac{\lambda_n ct}{a}\right) + B_n \cos\left(\frac{\lambda_n ct}{a}\right), \tag{7.3.56}$$

where λ_n satisfies the equation $J_0(\lambda) = 0$. Because $u(r,0) = 0$, and $T_n(0) = 0$, $B_n = 0$. Consequently, the product solution is

$$u(r,t) = \sum_{n=1}^{\infty} A_n J_0\left(\frac{\lambda_n r}{a}\right) \sin\left(\frac{\lambda_n ct}{a}\right). \tag{7.3.57}$$

To determine A_n, we use the condition

$$\frac{\partial u(r,0)}{\partial t} = \sum_{n=1}^{\infty} \frac{\lambda_n c}{a} A_n J_0\left(\frac{\lambda_n r}{a}\right) = \begin{cases} P/(\pi\epsilon^2\rho), & 0 \le r < \epsilon, \\ 0, & \epsilon < r < a. \end{cases} \tag{7.3.58}$$

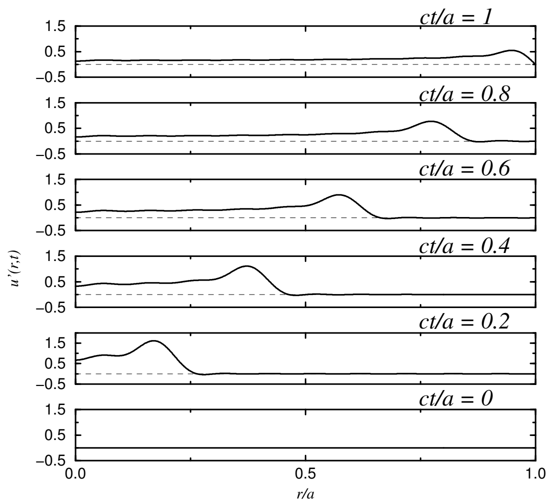


Figure 7.3.7: The axisymmetric vibrations $u'(r,t) = ca\rho u(r,t)/P$ of a circular membrane at various positions r/a at the times $ct/a = 0$, 0.2, 0.4, 0.6, 0.8, and 1 for $\epsilon = a/4$. Initially the membrane is struck by a hammer.

Equation 7.3.58 is a Fourier-Bessel expansion employing the orthogonal function $J_0(\lambda_n r/a)$, where

$$\frac{\lambda_n c}{a} A_n = \frac{2}{a^2 J_1^2(\lambda_n)} \int_0^\epsilon \frac{P}{\pi\epsilon^2\rho} J_0\left(\frac{\lambda_n r}{a}\right) r\,dr \tag{7.3.59}$$

from Equation 6.5.38 and Equation 6.5.45 in Section 6.5. Carrying out the integration,

$$A_n = \frac{2PJ_1(\lambda_n\epsilon/a)}{c\pi\epsilon\rho\lambda_n^2 J_1^2(\lambda_n)}, \tag{7.3.60}$$

or

$$u(r,t) = \frac{2P}{c\pi\epsilon\rho} \sum_{n=1}^{\infty} \frac{J_1(\lambda_n\epsilon/a)}{\lambda_n^2 J_1^2(\lambda_n)} J_0\left(\frac{\lambda_n r}{a}\right) \sin\left(\frac{\lambda_n ct}{a}\right). \tag{7.3.61}$$

Figures 7.3.7, 7.3.8, and 7.3.9 illustrate the solution, Equation 7.3.61, for various times and positions when $\epsilon = a/4$, and $\epsilon = a/20$. They were generated using the MATLAB script

```
% initialize parameters
clear; eps_over_a = 0.25; M = 20; dr = 0.02; dt = 0.02;
% load in zeros of J_0
zero( 1) =  2.40483; zero( 2) =  5.52008; zero( 3) =  8.65373;
zero( 4) = 11.79153; zero( 5) = 14.93092; zero( 6) = 18.07106;
zero( 7) = 21.21164; zero( 8) = 24.35247; zero( 9) = 27.49347;
zero(10) = 30.63461; zero(11) = 33.77582; zero(12) = 36.91710;
zero(13) = 40.05843; zero(14) = 43.19979; zero(15) = 46.34119;
zero(16) = 49.48261; zero(17) = 52.62405; zero(18) = 55.76551;
zero(19) = 58.90698; zero(20) = 62.04847;
% compute Fourier-Bessel coefficients
for m = 1:M
  a(m) = 2 * besselj(1,eps_over_a*zero(m)) ...
        / (eps_over_a*pi*zero(m)*zero(m)*besselj(1,zero(m))^2);
end
```

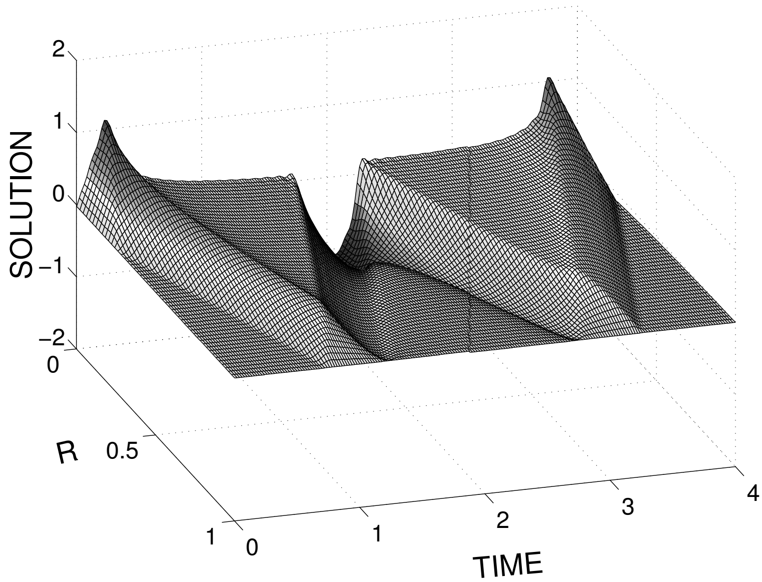


Figure 7.3.8: The axisymmetric vibrations $ca\rho u(r,t)/P$ of a circular membrane resulting from an initial hammer blow with $\epsilon = a/4$. The solution is plotted at various times ct/a and positions r/a.

```
R = [0:dr:1]; T = [0:dt:4];
u = zeros(length(T),length(R));
RR = repmat(R,[length(T) 1]);
TT = repmat(T',[1 length(R)]);
% compute solution from series solution
for m = 1:M
  u = u + a(m) .* besselj(0,zero(m)*RR) .* sin(zero(m)*TT);
end
% plot results
surf(RR,TT,u)
xlabel('R','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

Figures 7.3.8 and 7.3.9 show that striking the membrane with a hammer generates a pulse that propagates out to the rim, reflects, inverts, and propagates back to the center. This process then repeats forever.

Problems

Solve the wave equation $u_{tt} = c^2 u_{xx}$, $0 < x < L$, $0 < t$, subject to the boundary conditions that $u(0,t) = u(L,t) = 0, 0 < t$, and the following initial conditions for $0 < x < L$. Use MATLAB to illustrate your solution.

1. $u(x,0) = 0, \quad u_t(x,0) = 1$

2. $u(x,0) = 1, \quad u_t(x,0) = 0$

3. $u(x,0) = \begin{cases} 3hx/2L, & 0 < x < 2L/3, \\ 3h(L-x)/L, & 2L/3 < x < L, \end{cases} \quad u_t(x,0) = 0$

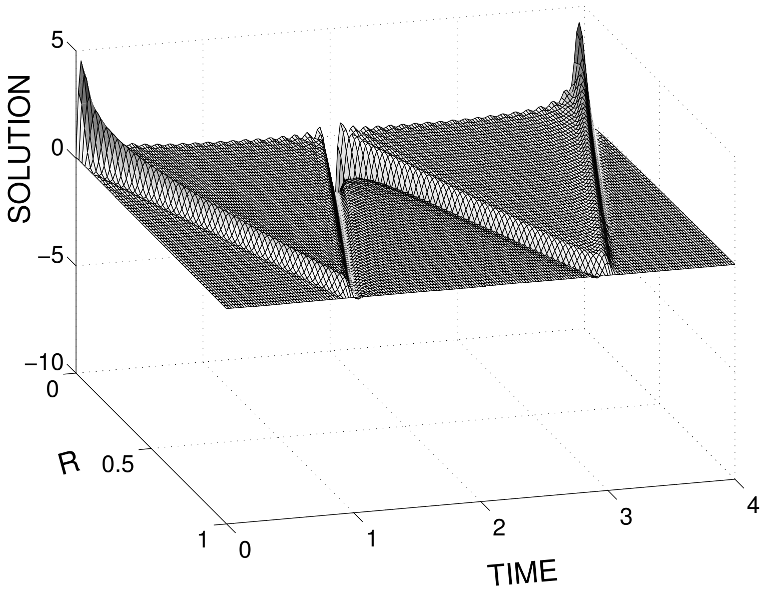


Figure 7.3.9: Same as Figure 7.3.8 except $\epsilon = a/20$.

4. $u(x,0) = [3\sin(\pi x/L) - \sin(3\pi x/L)]/4, \quad u_t(x,0) = 0,$

5. $u(x,0) = \sin(\pi x/L)\,, \quad u_t(x,0) = \begin{cases} 0, & 0 < x < L/4 \\ a, & L/4 < x < 3L/4 \\ 0, & 3L/4 < x < L \end{cases}$

6. $u(x,0) = 0, \quad u_t(x,0) = \begin{cases} ax/L, & 0 < x < L/2 \\ a(L-x)/L, & L/2 < x < L \end{cases}$

7. $u(x,0) = \begin{cases} x, & 0 < x < L/2, \\ L - x, & L/2 < x < L, \end{cases} \quad u_t(x,0) = 0$

8. Solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \pi, \qquad 0 < t,$$

subject to the boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = \frac{\partial u(\pi,t)}{\partial x} = 0, \qquad 0 < t,$$

and the initial conditions

$$u(x,0) = 0, \qquad \frac{\partial u(x,0)}{\partial t} = 1 + \cos^3(x), \qquad 0 < x < \pi.$$

Hint: You must include the separation constant of zero.

9. Solve[10] the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial x}\right), \qquad 0 \le x < 1, \qquad 0 < t,$$

subject to the boundary conditions

$$\lim_{x\to 0} |u(x,t)| < \infty, \qquad u(1,t) = 0, \qquad 0 < t,$$

and the initial conditions

$$u(x,0) = 0, \quad 0 \le x \le 1, \qquad \frac{\partial u(x,0)}{\partial t} = \begin{cases} 1, & 0 \le x < a, \\ 0, & a < x \le 1, \end{cases}$$

where $a < 1$. Hint: Use the substitution $4x = r^2$.

10. The differential equation for the longitudinal vibrations of a rod within a viscous fluid is

$$\frac{\partial^2 u}{\partial t^2} + 2h\frac{\partial u}{\partial t} = c^2\frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \qquad 0 < t,$$

where c is the velocity of sound in the rod and h is the damping coefficient. If the rod is fixed at $x = 0$ so that $u(0,t) = 0$, and allowed to freely oscillate at the other end $x = L$, so that $u_x(L,t) = 0$, find the vibrations for any location x and subsequent time t if the rod has the initial displacement of $u(x,0) = x$ and the initial velocity $u_t(x,0) = 0$ for $0 < x < L$. Assume that $h < c\pi/(2L)$. Why?

11. A closed pipe of length L contains air whose density is slightly greater than that of the outside air in the ratio of $1 + s_0$ to 1. Everything being at rest, we suddenly draw aside the disk closing one end of the pipe. We want to determine what happens *inside* the pipe after we remove the disk.

As the air rushes outside, it generates sound waves within the pipe. The wave equation

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

governs these waves, where c is the speed of sound and $u(x,t)$ is the velocity potential. Without going into the fluid mechanics of the problem, the boundary conditions are

a. No flow through the closed end: $u_x(0,t) = 0$.

b. No infinite acceleration at the open end: $u_{xx}(L,t) = 0$.

c. Air is initially at rest: $u_x(x,0) = 0$.

d. Air initially has a density greater than the surrounding air by the amount s_0: $u_t(x,0) = -c^2 s_0$.

Find the velocity potential at all positions within the pipe and all subsequent times.

12. One of the classic applications of the wave equation has been the explanation of the acoustic properties of string instruments. Usually we excite a string in one of three ways:

---

[10] Solved in a slightly different manner by Bailey, H., 2000: Motions of a hanging chain after the free end is given an initial velocity. *Am. J. Phys.*, **68**, 764–767.

by plucking (as in the harp, zither, etc.), by striking with a hammer (piano), or by bowing (violin, violoncello, etc.). In all of these cases, the governing partial differential equation is

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

with the boundary conditions $u(0,t) = u(L,t) = 0$, $0 < t$. For each of the following methods of exciting a string instrument, find the complete solution to the problem:

(a) Plucked string

For the initial conditions:

$$u(x,0) = \begin{cases} \beta x/a, & 0 < x < a, \\ \beta(L-x)/(L-a), & a < x < L, \end{cases}$$

and

$$u_t(x,0) = 0, \qquad 0 < x < L,$$

show that

$$u(x,t) = \frac{2\beta L^2}{\pi^2 a(L-a)} \sum_{n=1}^{\infty} \frac{1}{n^2} \sin\left(\frac{n\pi a}{L}\right) \sin\left(\frac{n\pi x}{L}\right) \cos\left(\frac{n\pi ct}{L}\right).$$

We note that the harmonics are absent where $\sin(n\pi a/L) = 0$. Thus, if we pluck the string at the center, all of the harmonics of even order are absent. Furthermore, the intensity of the successive harmonics varies as n^{-2}. The higher harmonics (overtones) are therefore relatively feeble compared to the $n = 1$ term (the fundamental).

(b) String excited by impact

The effect of the impact of a hammer depends upon the manner and duration of the contact, and is more difficult to estimate. However, as a first estimate, let

$$u(x,0) = 0, \qquad 0 < x < L,$$

and

$$u_t(x,0) = \begin{cases} \mu, & a - \epsilon < x < a + \epsilon, \\ 0, & \text{otherwise}, \end{cases}$$

where $\epsilon \ll 1$. Show that the solution in this case is

$$u(x,t) = \frac{4\mu L}{\pi^2 c} \sum_{n=1}^{\infty} \frac{1}{n^2} \sin\left(\frac{n\pi\epsilon}{L}\right) \sin\left(\frac{n\pi a}{L}\right) \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{n\pi ct}{L}\right).$$

As in part (a), the nth mode is absent if the origin is at a node. The intensity of the overtones are now of the same order of magnitude; higher harmonics (overtones) are relatively more in evidence than in part (a).

(c) Bowed violin string

The theory of the vibration of a string when excited by bowing is poorly understood. The bow drags the string for a time until the string springs back. After a while the process repeats. It can be shown[11] that the proper initial conditions are

$$u(x,0) = 0, \qquad 0 < x < L,$$

and

$$u_t(x,0) = 4\beta c(L-x)/L^2, \qquad 0 < x < L,$$

where β is the maximum displacement. Show that the solution is now

$$u(x,t) = \frac{8\beta}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{n\pi ct}{L}\right).$$

7.4 D'ALEMBERT'S FORMULA

In the previous section we sought solutions to the homogeneous wave equation in the form of a product $X(x)T(t)$. For the one-dimensional wave equation there is a more general method for constructing the solution, published by d'Alembert[12] in 1747.

Let us determine a solution to the homogeneous wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < \infty, \quad 0 < t, \tag{7.4.1}$$

which satisfies the initial conditions

$$u(x,0) = f(x), \quad \frac{\partial u(x,0)}{\partial t} = g(x), \quad -\infty < x < \infty. \tag{7.4.2}$$

We begin by introducing two new variables ξ, η defined by $\xi = x + ct$, and $\eta = x - ct$, and set $u(x,t) = w(\xi,\eta)$. The variables ξ and η are called the *characteristics* of the wave equation. Using the chain rule,

$$\frac{\partial}{\partial x} = \frac{\partial \xi}{\partial x}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial x}\frac{\partial}{\partial \eta} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \tag{7.4.3}$$

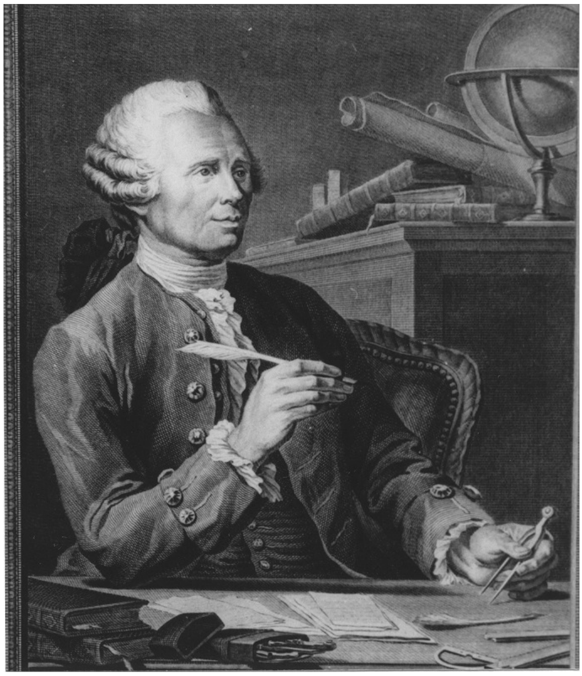
$$\frac{\partial}{\partial t} = \frac{\partial \xi}{\partial t}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial t}\frac{\partial}{\partial \eta} = c\frac{\partial}{\partial \xi} - c\frac{\partial}{\partial \eta} \tag{7.4.4}$$

$$\frac{\partial^2}{\partial x^2} = \frac{\partial \xi}{\partial x}\frac{\partial}{\partial \xi}\left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right) + \frac{\partial \eta}{\partial x}\frac{\partial}{\partial \eta}\left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right) \tag{7.4.5}$$

$$= \frac{\partial^2}{\partial \xi^2} + 2\frac{\partial^2}{\partial \xi \partial \eta} + \frac{\partial^2}{\partial \eta^2}, \tag{7.4.6}$$

[11] See Lamb, H., 1960: *The Dynamical Theory of Sound*. Dover Publishers, Section 27.

[12] D'Alembert, J., 1747: Recherches sur la courbe que forme une corde tenduë mise en vibration. *Hist. Acad. R. Sci. Belles Lett., Berlin*, 214–219.



Although largely self-educated in mathematics, Jean Le Rond d'Alembert (1717–1783) gained equal fame as a mathematician and *philosophe* of the continental Enlightenment. By the middle of the eighteenth century, he stood with such leading European mathematicians and mathematical physicists as Clairaut, D. Bernoulli, and Euler. Today we best remember him for his work in fluid dynamics and applying partial differential equations to problems in physics. (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

and similarly

$$\frac{\partial^2}{\partial t^2} = c^2\left(\frac{\partial^2}{\partial \xi^2} - 2\frac{\partial^2}{\partial \xi \partial \eta} + \frac{\partial^2}{\partial \eta^2}\right), \tag{7.4.7}$$

so that the wave equation becomes

$$\frac{\partial^2 w}{\partial \xi \partial \eta} = 0. \tag{7.4.8}$$

The general solution of Equation 7.4.8 is

$$w(\xi, \eta) = F(\xi) + G(\eta). \tag{7.4.9}$$

Thus, the general solution of Equation 7.4.1 is of the form

$$u(x,t) = F(x+ct) + G(x-ct), \tag{7.4.10}$$

where F and G are arbitrary functions of one variable and are assumed to be twice differentiable. Setting $t = 0$ in Equation 7.4.10 and using the initial condition that $u(x,0) = f(x)$,

$$F(x) + G(x) = f(x). \tag{7.4.11}$$

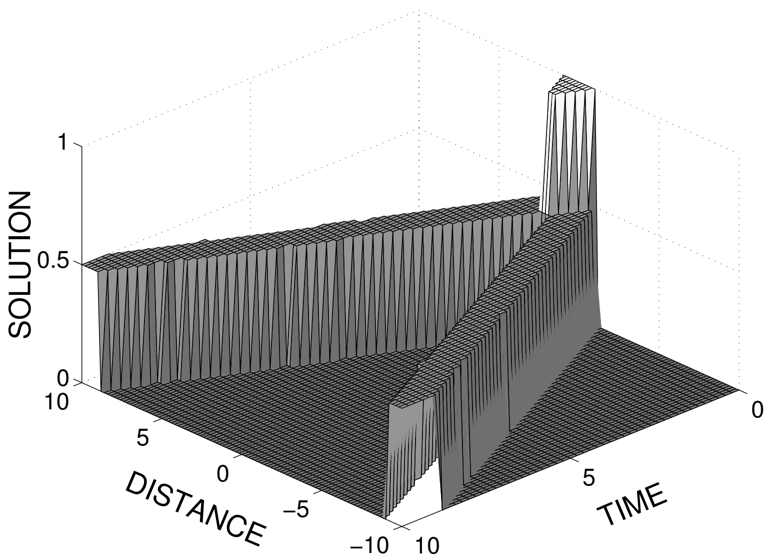


Figure 7.4.1: D'Alembert's solution, Equation 7.4.18, to the wave equation.

The partial derivative of Equation 7.4.10 with respect to t yields

$$\frac{\partial u(x,t)}{\partial t} = cF'(x+ct) - cG'(x-ct). \tag{7.4.12}$$

Here primes denote differentiation with respect to the argument of the function. If we set $t=0$ in Equation 7.4.12 and apply the initial condition that $u_t(x,0)=g(x)$,

$$cF'(x) - cG'(x) = g(x). \tag{7.4.13}$$

Integrating Equation 7.4.13 from 0 to any point x gives

$$F(x) - G(x) = \frac{1}{c}\int_0^x g(\tau)\,d\tau + C, \tag{7.4.14}$$

where C is the constant of integration. Combining this result with Equation 7.4.11,

$$F(x) = \frac{f(x)}{2} + \frac{1}{2c}\int_0^x g(\tau)\,d\tau + \frac{C}{2}, \tag{7.4.15}$$

and

$$G(x) = \frac{g(x)}{2} - \frac{1}{2c}\int_0^x g(\tau)\,d\tau - \frac{C}{2}. \tag{7.4.16}$$

If we replace the variable x in the expression for F and G by $x+ct$ and $x-ct$, respectively, and substitute the results into Equation 7.4.10, we finally arrive at the formula

$$u(x,t) = \frac{f(x+ct)+f(x-ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\tau)\,d\tau. \tag{7.4.17}$$

This is known as *d'Alembert's formula* for the solution of the wave equation, Equation 7.4.1, subject to the initial conditions, Equation 7.4.2. It gives a *representation* of the solution in terms of *known* initial conditions.

• Example 7.4.1

To illustrate d'Alembert's formula, let us find the solution to the wave equation, Equation 7.4.1, satisfying the initial conditions $u(x,0) = H(x+1) - H(x-1)$ and $u_t(x,0) = 0$, $-\infty < x < \infty$. By d'Alembert's formula, Equation 7.4.17,

$$u(x,t) = \tfrac{1}{2}\left[H(x+ct+1) + H(x-ct+1) - H(x+ct-1) - H(x-ct-1)\right]. \quad \textbf{(7.4.18)}$$

We illustrate this solution in Figure 7.4.1 generated by the MATLAB script

```
% set mesh size for solution
clear; dx = 0.1; dt = 0.1;
% compute grid
X=[-10:dx:10]; T = [0:dt:10];
for j=1:length(T); t = T(j);
for i=1:length(X); x = X(i);
% compute characteristics
  characteristic_1 = x + t; characteristic_2 = x - t;
% compute solution
  XX(i,j) = x; TT(i,j) = t;
  u(i,j) = 0.5*(stepfun(characteristic_1,-1)+stepfun(characteristic_2,-1)...
                -stepfun(characteristic_1, 1)-stepfun(characteristic_2, 1));
end; end
surf(XX,TT,u); colormap autumn;
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

In this figure, you can clearly see the characteristics as they emanate from the discontinuities at $x = \pm 1$. □

• Example 7.4.2

Let us find the solution to the wave equation, Equation 7.4.1, when $u(x,0) = 0$, and $u_t(x,0) = \sin(2x)$, $-\infty < x < \infty$. By d'Alembert's formula, the solution is

$$u(x,t) = \frac{1}{2c}\int_{x-ct}^{x+ct} \sin(2\tau)\,d\tau = \frac{\sin(2x)\sin(2ct)}{2}. \quad \textbf{(7.4.19)}$$

In addition to providing a method of solving the wave equation, d'Alembert's solution can also provide physical insight into the vibration of a string. Consider the case when we release a string with zero velocity after giving it an initial displacement of $f(x)$. According to Equation 7.4.17, the displacement at a point x at any time t is

$$u(x,t) = \frac{f(x+ct) + f(x-ct)}{2}. \quad \textbf{(7.4.20)}$$

Because the function $f(x-ct)$ is the same as the function of $f(x)$ translated to the right by a distance equal to ct, $f(x-ct)$ represents a wave of form $f(x)$ traveling to the right with the velocity c, a forward wave. Similarly, we can interpret the function $f(x+ct)$ as representing a wave with the shape $f(x)$ traveling to the left with the velocity c, a backward wave. Thus, the solution, Equation 7.4.17, is a superposition of forward and backward waves traveling with the same velocity c and having the shape of the initial profile $f(x)$ with half of the

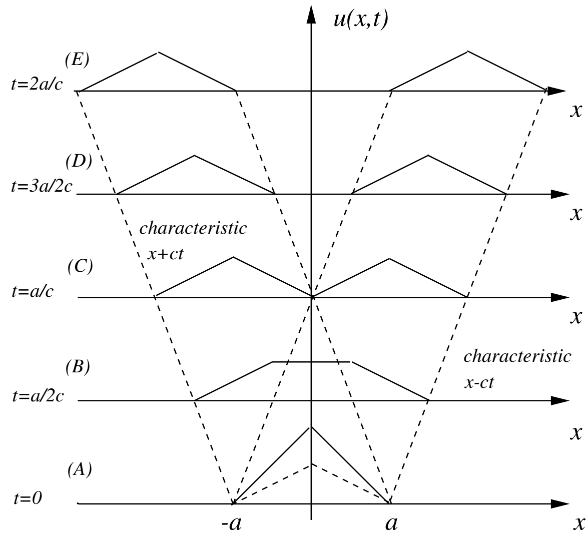


Figure 7.4.2: The propagation of waves due to an initial displacement according to d'Alembert's formula.

amplitude. Clearly the characteristics $x + ct$ and $x - ct$ give the propagation paths along which the waveform $f(x)$ propagates. □

• Example 7.4.3

To illustrate our physical interpretation of d'Alembert's solution, suppose that the string has an initial displacement defined by

$$f(x) = \begin{cases} a - |x|, & -a \le x \le a, \\ 0, & \text{otherwise.} \end{cases} \tag{7.4.21}$$

In Figure 7.4.2(A) the forward and backward waves, indicated by the dashed line, coincide at $t = 0$. As time advances, both waves move in opposite directions. In particular, at $t = a/(2c)$, they moved through a distance $a/2$, resulting in the displacement of the string shown in Figure 7.4.2(B). Eventually, at $t = a/c$, the forward and backward waves completely separate. Finally, Figures 7.4.2(D) and 7.4.3(E) show how the waves radiate off to infinity at the speed of c. Note that at each point the string returns to its original position of rest after the passage of each wave.

Consider now the opposite situation when $u(x, 0) = 0$, and $u_t(x, 0) = g(x)$. The displacement is

$$u(x,t) = \frac{1}{2c} \int_{x-ct}^{x+ct} g(\tau)\, d\tau. \tag{7.4.22}$$

If we introduce the function

$$\varphi(x) = \frac{1}{2c} \int_0^x g(\tau)\, d\tau, \tag{7.4.23}$$

then we can write Equation 7.4.22 as

$$u(x,t) = \varphi(x+ct) - \varphi(x-ct), \tag{7.4.24}$$

which again shows that the solution is a superposition of a forward wave $-\varphi(x - ct)$ and a backward wave $\varphi(x + ct)$ traveling with the same velocity c. The function φ, which we

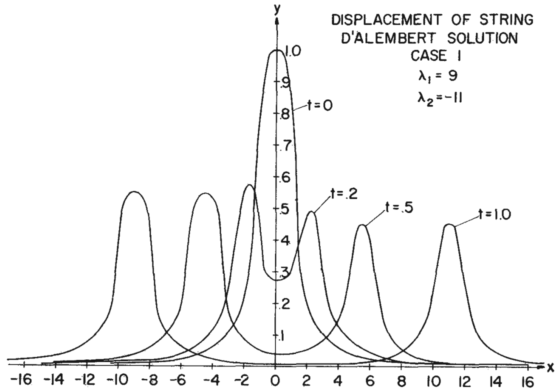


Figure 7.4.3: Displacement of an infinite, moving threadline when $c = 10$, and $V = 1$.

compute from Equation 7.4.23 and the initial velocity $g(x)$, determines the exact form of these waves. □

• Example 7.4.4: Vibration of a moving threadline

The characterization and analysis of the oscillations of a string or yarn have an important application in the textile industry because they describe the way that yarn winds on a bobbin.[13] As we showed in Section 7.1, the governing equation, the "threadline equation," is

$$\frac{\partial^2 u}{\partial t^2} + \alpha \frac{\partial^2 u}{\partial x \partial t} + \beta \frac{\partial^2 u}{\partial x^2} = 0, \tag{7.4.25}$$

where $\alpha = 2V$, $\beta = V^2 - gT/\rho$, V is the windup velocity, g is the gravitational attraction, T is the tension in the yarn, and ρ is the density of the yarn. We now introduce the characteristics $\xi = x + \lambda_1 t$, and $\eta = x + \lambda_2 t$, where λ_1 and λ_2 are yet undetermined. Upon substituting ξ and η into Equation 7.4.25,

$$(\lambda_1^2 + 2V\lambda_1 + V^2 - gT/\rho)u_{\xi\xi} + (\lambda_2^2 + 2V\lambda_2 + V^2 - gT/\rho)u_{\eta\eta} \\ + [2V^2 - 2gT/\rho + 2V(\lambda_1 + \lambda_2) + 2\lambda_1\lambda_2]u_{\xi\eta} = 0. \tag{7.4.26}$$

If we choose λ_1 and λ_2 to be roots of the equation

$$\lambda^2 + 2V\lambda + V^2 - gT/\rho = 0, \tag{7.4.27}$$

Equation 7.4.26 reduces to the simple form

$$u_{\xi\eta} = 0, \tag{7.4.28}$$

which has the general solution

$$u(x,t) = F(\xi) + G(\eta) = F(x + \lambda_1 t) + G(x + \lambda_2 t). \tag{7.4.29}$$

Solving Equation 7.4.27 yields

$$\lambda_1 = c - V, \qquad \text{and} \qquad \lambda_2 = -c - V, \tag{7.4.30}$$

[13] See Swope, R. D., and W. F. Ames, 1963: Vibrations of a moving threadline. *J. Franklin Inst.*, **275**, 36–55.

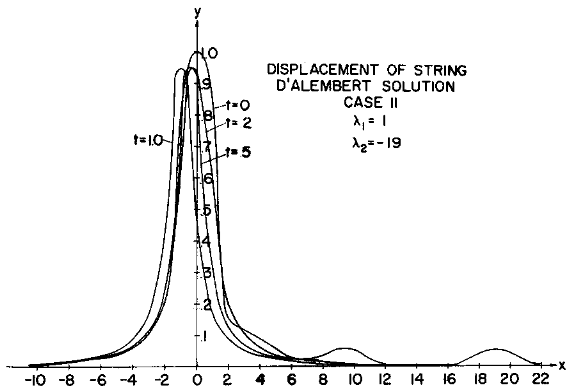


Figure 7.4.4: Displacement of an infinite, moving threadline when $c = 11$, and $V = 10$.

where $c = \sqrt{gT/\rho}$. If the initial conditions are

$$u(x, 0) = f(x), \qquad \text{and} \qquad u_t(x, 0) = g(x), \tag{7.4.31}$$

then

$$u(x,t) = \frac{1}{2c}\left[\lambda_1 f(x+\lambda_2 t) - \lambda_2 f(x+\lambda_1 t) + \int_{x+\lambda_2 t}^{x+\lambda_1 t} g(\tau)\, d\tau\right]. \tag{7.4.32}$$

Because λ_1 does not generally equal to λ_2, the two waves that constitute the motion of the string move with different speeds and have different shapes and forms. For example, if

$$f(x) = \frac{1}{x^2+1}, \qquad \text{and} \qquad g(x) = 0, \tag{7.4.33}$$

$$u(x,t) = \frac{1}{2c}\left\{\frac{c-V}{1+[x-(c+V)t]^2} + \frac{c+V}{1+[x-(c-V)t]^2}\right\}. \tag{7.4.34}$$

Figures 7.4.3 and 7.4.4 illustrate this solution for several different parameters.

Problems

Use d'Alembert's formula to solve the wave equation, Equation 7.4.1, for the following initial conditions defined for $|x| < \infty$. Then illustrate your solution using MATLAB.

1. $u(x, 0) = 2\sin(x)\cos(x)$ $\qquad u_t(x, 0) = \cos(x)$
2. $u(x, 0) = x\sin(x)$ $\qquad u_t(x, 0) = \cos(2x)$
3. $u(x, 0) = 1/(x^2 + 1)$ $\qquad u_t(x, 0) = e^x$
4. $u(x, 0) = e^{-x}$ $\qquad u_t(x, 0) = 1/(x^2 + 1)$
5. $u(x, 0) = \cos(\pi x/2)$ $\qquad u_t(x, 0) = \sinh(ax)$
6. $u(x, 0) = \sin(3x)$ $\qquad u_t(x, 0) = \sin(2x) - \sin(x)$

7. Assuming that the functions F and G are differentiable, show by direct substitution that

$$u(x,t) = EF(x+ct) - EG(x-ct) - \tfrac{1}{8}kc^2t^2 + \tfrac{3}{8}kx^2,$$

and

$$v(x,t) = cF(x+ct) + cG(x-ct) - \frac{kc^2xt}{4E}$$

are the d'Alembert solutions to the hyperbolic system

$$\frac{\partial u}{\partial t} = E\frac{\partial v}{\partial x}, \qquad \frac{\partial u}{\partial x} = \rho\frac{\partial v}{\partial t} + kx, \qquad -\infty < x < \infty, \quad 0 < t,$$

where $c^2 = E/\rho$ and E, k, and ρ are constants.

8. D'Alembert's solution can also be used in problems over the limited domain $0 < x < L$. To illustrate this, let us solve the wave equation, Equation 7.4.1, with the initial conditions $u(x,0) = 0$, $u_t(x,0) = V_{max}(1 - x/L)$, $0 < x < L$, and the boundary conditions $u(0,t) = u(L,t) = 0$, $0 < t$.

Step 1: Show that the solution to this problem is

$$u(x,t) = \tfrac{1}{2}[V_0(x+ct) - V_0(x-ct)],$$

where

$$V_0(\chi) = \frac{1}{c}\int_0^{\chi} u_t(\xi,0)\,d\xi = \frac{V_{max}\chi}{c}\left(1 - \frac{\chi}{2L}\right), \qquad 0 < \chi < L,$$

along with the periodicity conditions $V_0(\chi) = V_0(-\chi)$, and $V_0(L+\chi) = V_0(L-\chi)$ to take care of those cases when the argument of $V_0(\cdot)$ is outside of $(0, L)$. Hint: Substitute the solution into the boundary conditions.

Step 2: Show that at any point x within the interval $(0, L)$, the solution repeats with a period of $2L/c$ if $ct > 2L$. Therefore, if we know the behavior of the solution for the time interval $0 < ct < 2L$, we know the behavior for any other time.

Step 3: Show that the solution at any point x within the interval $(0, L)$ and time $t + L/c$, where $0 < ct < L$, is the mirror image (about $u = 0$) of the solution at the point $L - x$ and time t, where $0 < ct < L$.

Step 4: Show that the maximum value of $u(x,t)$ occurs at $x = ct$, where $0 < x < L$ and when $0 < ct < L$. At that point,

$$u_{max} = \frac{V_{max}x}{c}\left(1 - \frac{x}{L}\right),$$

where u_{max} equals the largest magnitude of $u(x,t)$ for any time t. Plot u_{max} as a function x and show that it is a parabola. Hint: Find the maximum value of $u(x,t)$ when $0 < x \le ct$ and $ct \le x < L$ with $0 < x + ct < L$ or $L < x + ct < 2L$.

7.5 NUMERICAL SOLUTION OF THE WAVE EQUATION

Despite the powerful techniques shown in the previous sections for solving the wave equation, often these analytic techniques fail and we must resort to numerical techniques. In contrast to the continuous solutions, finite difference methods, a type of numerical solution technique, give discrete numerical values at a specific location (x_m, t_n), called a *grid point*. These numerical values represent a numerical approximation of the continuous solution over the region $(x_m - \Delta x/2, x_m + \Delta x/2)$ and $(t_n - \Delta t/2, t_n + \Delta t/2)$, where Δx and Δt are the distance and time intervals between grid points, respectively. Clearly, in the limit of

$\Delta x, \Delta t \to 0$, we recover the continuous solution. However, practical considerations such as computer memory or execution time often require that Δx and Δt, although small, are not negligibly small.

The first task in the numerical solution of a partial differential equation is the replacement of its continuous derivatives with finite differences. The most popular approach employs Taylor expansions. If we focus on the x-derivative, then the value of the solution at $u[(m+1)\Delta x, n\Delta t]$ in terms of the solution at $(m\Delta x, n\Delta t)$ is

$$\begin{aligned} u[(m+1)\Delta x, n\Delta t] &= u(x_m, t_n) + \frac{\Delta x}{1!}\frac{\partial u(x_m,t_n)}{\partial x} + \frac{(\Delta x)^2}{2!}\frac{\partial^2 u(x_m,t_n)}{\partial x^2} \\ &\quad + \frac{(\Delta x)^3}{3!}\frac{\partial^3 u(x_m,t_n)}{\partial x^3} + \frac{(\Delta x)^4}{4!}\frac{\partial^4 u(x_m,t_n)}{\partial x^4} + \cdots \qquad \textbf{(7.5.1)} \\ &= u(x_m,t_n) + \Delta x\frac{\partial u(x_m,t_n)}{\partial x} + O[(\Delta x)^2], \qquad \textbf{(7.5.2)} \end{aligned}$$

where $O[(\Delta x)^2]$ gives a measure of the magnitude of neglected terms.[14]

From Equation 7.5.2, one possible approximation for u_x is

$$\frac{\partial u(x_m,t_n)}{\partial x} = \frac{u_{m+1}^n - u_m^n}{\Delta x} + O(\Delta x), \qquad \textbf{(7.5.3)}$$

where we use the standard notation that $u_m^n = u(x_m, t_n)$. This is an example of a *one-sided finite difference* approximation of the partial derivative u_x. The error in using this approximation grows as Δx.

Another possible approximation for the derivative arises from using $u(m\Delta x, n\Delta t)$ and $u[(m-1)\Delta x, n\Delta t]$. From the Taylor expansion:

$$\begin{aligned} u[(m-1)\Delta x, n\Delta t] &= u(x_m, t_n) - \frac{\Delta x}{1!}\frac{\partial u(x_m,t_n)}{\partial x} + \frac{(\Delta x)^2}{2!}\frac{\partial^2 u(x_m,t_n)}{\partial x^2} \\ &\quad - \frac{(\Delta x)^3}{3!}\frac{\partial^3 u(x_m,t_n)}{\partial x^3} + \frac{(\Delta x)^4}{4!}\frac{\partial^4 u(x_m,t_n)}{\partial x^4} - \cdots, \qquad \textbf{(7.5.4)} \end{aligned}$$

we can also obtain the one-sided difference formula

$$\frac{u(x_m,t_n)}{\partial x} = \frac{u_m^n - u_{m-1}^n}{\Delta x} + O(\Delta x). \qquad \textbf{(7.5.5)}$$

A third possibility arises from subtracting Equation 7.5.4 from Equation 7.5.1:

$$u_{m+1}^n - u_{m-1}^n = 2\Delta x\frac{\partial u(x_m,t_n)}{\partial x} + O[(\Delta x)^3], \qquad \textbf{(7.5.6)}$$

or

$$\frac{\partial u(x_m,t_n)}{\partial x} = \frac{u_{m+1}^n - u_{m-1}^n}{2\Delta x} + O[(\Delta x)^2]. \qquad \textbf{(7.5.7)}$$

Thus, the choice of the finite differencing scheme can produce profound differences in the accuracy of the results. In the present case, *centered finite differences* can yield results that are markedly better than using one-sided differences.

---

[14] The symbol O is a mathematical notation indicating relative magnitude of terms, namely that $f(\epsilon) = O(\epsilon^n)$ provided $\lim_{\epsilon\to 0}|f(\epsilon)/\epsilon^n| < \infty$. For example, as $\epsilon \to 0$, $\sin(\epsilon) = O(\epsilon)$, $\sin(\epsilon^2) = O(\epsilon^2)$, and $\cos(\epsilon) = O(1)$.

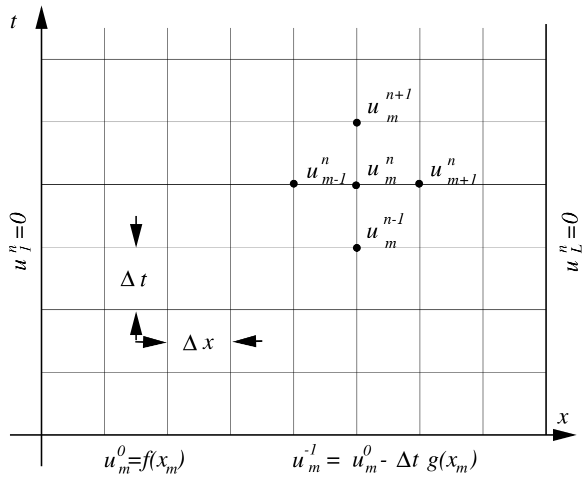


Figure 7.5.1: Schematic of the numerical solution of the wave equation with fixed endpoints.

To solve the wave equation, we need to approximate u_{xx}. If we add Equation 7.5.1 and Equation 7.5.4,

$$u^n_{m+1} + u^n_{m-1} = 2u^n_m + \frac{\partial^2 u(x_m, t_n)}{\partial x^2}(\Delta x)^2 + O[(\Delta x)^4], \tag{7.5.8}$$

or

$$\frac{\partial^2 u(x_m, t_n)}{\partial x^2} = \frac{u^n_{m+1} - 2u^n_m + u^n_{m-1}}{(\Delta x)^2} + O[(\Delta x)^2]. \tag{7.5.9}$$

Similar considerations hold for the time derivative. Thus, by neglecting errors of $O[(\Delta x)^2]$ and $O[(\Delta t)^2]$, we may approximate the wave equation by

$$\frac{u^{n+1}_m - 2u^n_m + u^{n-1}_m}{(\Delta t)^2} = c^2 \frac{u^n_{m+1} - 2u^n_m + u^n_{m-1}}{(\Delta x)^2}. \tag{7.5.10}$$

Because the wave equation represents evolutionary change of some quantity, Equation 7.5.10 is generally used as a predictive equation where we forecast u^{n+1}_m by

$$u^{n+1}_m = 2u^n_m - u^{n-1}_m + \left(\frac{c\Delta t}{\Delta x}\right)^2 \left(u^n_{m+1} - 2u^n_m + u^n_{m-1}\right). \tag{7.5.11}$$

Figure 7.5.1 illustrates this numerical scheme.

The greatest challenge in using Equation 7.5.11 occurs with the very first prediction. When $n = 0$, clearly u^0_{m+1}, u^0_m, and u^0_{m-1} are specified from the initial condition $u(m\Delta x, 0) = f(x_m)$. But what about u^{-1}_m? Recall that we still have $u_t(x, 0) = g(x)$. If we use the backward difference formula, Equation 7.5.5,

$$\frac{u^0_m - u^{-1}_m}{\Delta t} = g(x_m). \tag{7.5.12}$$

Solving for u^{-1}_m,

$$u^{-1}_m = u^0_m - \Delta t g(x_m). \tag{7.5.13}$$

One disadvantage of using the backward finite-difference formula is the larger error associated with this term compared to those associated with the finite-differenced form of the wave equation. In the case of the barotropic vorticity equation, a partial differential equation with wave-like solutions, this inconsistency eventually leads to a separation of solution between adjacent time levels.[15] This difficulty is avoided by stopping after a certain number of time steps, averaging the solution, and starting again.

A better solution for computing that first time step employs the centered difference form

$$\frac{u_m^1 - u_m^{-1}}{2\Delta t} = g(x_m), \tag{7.5.14}$$

along with the wave equation

$$\frac{u_m^1 - 2u_m^0 + u_m^{-1}}{(\Delta t)^2} = c^2 \frac{u_{m+1}^0 - 2u_m^0 + u_{m-1}^0}{(\Delta x)^2}, \tag{7.5.15}$$

so that

$$u_m^1 = \left(\frac{c\Delta t}{\Delta x}\right)^2 \frac{f(x_{m+1}) + f(x_{m-1})}{2} + \left[1 - \left(\frac{c\Delta t}{\Delta x}\right)^2\right] f(x_m) + \Delta t g(x_m). \tag{7.5.16}$$

Although it appears that we are ready to start calculating, we need to check whether our numerical scheme possesses three properties: convergence, stability, and consistency. By *consistency* we mean that the difference equations approach the differential equation as $\Delta x, \Delta t \to 0$. To prove consistency, we first write u_{m+1}^n, u_{m-1}^n, u_m^{n-1}, and u_m^{n+1} in terms of $u(x,t)$ and its derivatives evaluated at (x_m, t_n). From Taylor expansions,

$$u_{m+1}^n = u_m^n + \Delta x \left.\frac{\partial u}{\partial x}\right|_n^m + \tfrac{1}{2}(\Delta x)^2 \left.\frac{\partial^2 u}{\partial x^2}\right|_n^m + \tfrac{1}{6}(\Delta x)^3 \left.\frac{\partial^3 u}{\partial x^3}\right|_n^m + \cdots, \tag{7.5.17}$$

$$u_{m-1}^n = u_m^n - \Delta x \left.\frac{\partial u}{\partial x}\right|_n^m + \tfrac{1}{2}(\Delta x)^2 \left.\frac{\partial^2 u}{\partial x^2}\right|_n^m - \tfrac{1}{6}(\Delta x)^3 \left.\frac{\partial^3 u}{\partial x^3}\right|_n^m + \cdots, \tag{7.5.18}$$

$$u_m^{n+1} = u_m^n + \Delta t \left.\frac{\partial u}{\partial t}\right|_n^m + \tfrac{1}{2}(\Delta t)^2 \left.\frac{\partial^2 u}{\partial t^2}\right|_n^m + \tfrac{1}{6}(\Delta t)^3 \left.\frac{\partial^3 u}{\partial t^3}\right|_n^m + \cdots, \tag{7.5.19}$$

and

$$u_m^{n-1} = u_m^n - \Delta t \left.\frac{\partial u}{\partial t}\right|_n^m + \tfrac{1}{2}(\Delta t)^2 \left.\frac{\partial^2 u}{\partial t^2}\right|_n^m - \tfrac{1}{6}(\Delta t)^3 \left.\frac{\partial^3 u}{\partial t^3}\right|_n^m + \cdots. \tag{7.5.20}$$

Substituting Equations 7.5.17 through 7.5.20 into Equation 7.5.10, we obtain

$$\begin{aligned} \frac{u_m^{n+1} - 2u_m^n + u_m^{n-1}}{(\Delta t)^2} &- c^2 \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{(\Delta x)^2} \\ &= \left.\left(\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2}\right)\right|_n^m + \tfrac{1}{12}(\Delta t)^2 \left.\frac{\partial^4 u}{\partial t^4}\right|_n^m - \tfrac{1}{12}(c\Delta x)^2 \left.\frac{\partial^4 u}{\partial x^4}\right|_n^m + \cdots. \end{aligned} \tag{7.5.21}$$

The first term on the right side of Equation 7.5.21 vanishes because $u(x,t)$ satisfies the wave equation. As $\Delta x \to 0$, $\Delta t \to 0$, the remaining terms on the right side of Equation 7.5.21

[15] Gates, W. L., 1959: On the truncation error, stability, and convergence of difference solutions of the barotropic vorticity equation. *J. Meteorol.*, **16**, 556–568. See Section 4.

tend to zero and Equation 7.5.10 is a consistent finite difference approximation of the wave equation.

Stability is another question. Under certain conditions the small errors inherent in fixed precision arithmetic (round off) can grow for certain choices of Δx and Δt. During the 1920s the mathematicians Courant, Friedrichs, and Lewy[16] found that if $c\Delta t/\Delta x > 1$, then our scheme is unstable. This CFL criterion has its origin in the fact that if $c\Delta t > \Delta x$, then we are asking signals in the numerical scheme to travel faster than their real-world counterparts and this unrealistic expectation leads to instability!

One method of determining *stability*, commonly called the von Neumann method,[17] involves examining solutions to Equation 7.5.11 that have the form

$$u_m^n = e^{im\theta} e^{in\lambda}, \tag{7.5.22}$$

where θ is an arbitrary real number and λ is a yet undetermined complex number. Our choice of Equation 7.5.22 is motivated by the fact that the initial condition u_m^0 can be represented by a Fourier series where a typical term behaves as $e^{im\theta}$.

If we substitute Equation 7.5.22 into Equation 7.5.10 and divide out the common factor $e^{im\theta}e^{in\lambda}$, we have that

$$\frac{e^{i\lambda} - 2 + e^{-i\lambda}}{(\Delta t)^2} = c^2 \frac{e^{i\theta} - 2 + e^{-i\theta}}{(\Delta x)^2}, \tag{7.5.23}$$

or

$$\sin^2\left(\frac{\lambda}{2}\right) = \left(\frac{c\Delta t}{\Delta x}\right)^2 \sin^2\left(\frac{\theta}{2}\right). \tag{7.5.24}$$

The behavior of u_m^n is determined by the values of λ given by Equation 7.5.24. If $c\Delta t/\Delta x \le 1$, then λ is real and u_m^n is bounded for all θ as $n \to \infty$. If $c\Delta t/\Delta x > 1$, then it is possible to find a value of θ such that the right side of Equation 7.5.24 exceeds unity and the corresponding values of λ occur as complex conjugate pairs. The λ with the negative imaginary part produces a solution with exponential growth because $n = t_n/\Delta t \to \infty$ as $\Delta t \to 0$ for a fixed t_n and $c\Delta t/\Delta x$. Thus, the value of u_m^n becomes infinitely large, even though the initial data may be arbitrarily small.

Finally, we must check for convergence. A numerical scheme is *convergent* if the numerical solution approaches the continuous solution as $\Delta x, \Delta t \to 0$. The general procedure for proving convergence involves the evolution of the error term e_m^n, which gives the difference between the true solution $u(x_m, t_n)$ and the finite difference solution u_m^n. From Equation 7.5.21,

$$e_m^{n+1} = \left(\frac{c\Delta t}{\Delta x}\right)^2 \left(e_{m+1}^n + e_{m-1}^n\right) + 2\left[1 - \left(\frac{c\Delta t}{\Delta x}\right)^2\right] e_m^n - e_m^{n-1} + O[(\Delta t)^4] + O[(\Delta x)^2(\Delta t)^2]. \tag{7.5.25}$$

Let us apply Equation 7.5.25 to work backwards from the point (x_m, t_n) by changing n to $n-1$. The nonvanishing terms in e_m^n reduce to a sum of $n+1$ values on the line $n=1$ plus $\frac{1}{2}(n+1)n$ terms of the form $A(\Delta x)^4$. If we define the max norm $||e_n|| = \max_m |e_m^n|$, then

$$||e_n|| \le nB(\Delta x)^3 + \tfrac{1}{2}(n+1)nA(\Delta x)^4. \tag{7.5.26}$$

---

[16] Courant, R., K. O. Friedrichs, and H. Lewy, 1928: Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Annalen*, **100**, 32–74. Translated into English in *IBM J. Res. Dev.*, **11**, 215–234.

[17] After its inventor, J. von Neumann. See O'Brien, G. G., M. A. Hyman, and S. Kaplan, 1950: A study of the numerical solution of partial differential equations. *J. Math. Phys. (Cambridge, MA)*, **29**, 223–251.

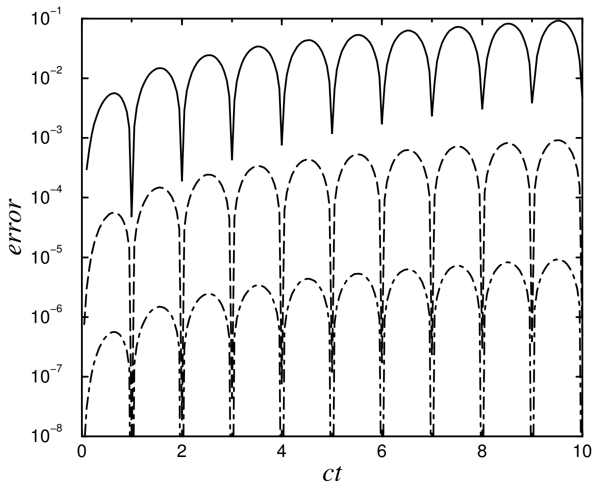


Figure 7.5.2: The growth of error $||e_n||$ as a function of ct for various resolutions. For the top line, $\Delta x = 0.1$; for the middle line, $\Delta x = 0.01$; and for the bottom line, $\Delta x = 0.001$.

Because $n\Delta x \le ct_n$, Equation 7.5.26 simplifies to

$$||e_n|| \le ct_n B(\Delta x)^2 + \tfrac{1}{2}c^2 t_n^2 A(\Delta x)^2. \qquad \textbf{(7.5.27)}$$

Thus, the error tends to zero as $\Delta x \to 0$, verifying convergence. We illustrate Equation 7.5.27 by using the finite difference equation, Equation 7.5.11, to compute $||e_n||$ during a numerical experiment that used $c\Delta t/\Delta x = 0.5$, $f(x) = \sin(\pi x)$, and $g(x) = 0$; $||e_n||$ is plotted in Figure 7.5.2. Note how each increase of resolution by 10 results in a drop in the error by 100.

In the following examples we apply our scheme to solve a few simple initial and boundary conditions:

• Example 7.5.1

For our first example, we resolve Equation 7.3.1 through Equation 7.3.3 and Equation 7.3.25 and Equation 7.3.26 numerically using MATLAB. The MATLAB code is

```
clear
coeff = 0.5; coeffsq = coeff * coeff % coeff = cΔt/Δx
dx = 0.04; dt = coeff * dx; N = 100; x = 0:dx:1;
M = 1/dx + 1; % M = number of spatial grid points
% introduce the initial conditions via F and G
F = zeros(M,1); G = zeros(M,1);
for m = 1:M
  if x(m) >= 0.25 & x(m) <= 0.5
      F(m) = 4 * x(m) - 1; end
  if x(m) >= 0.5 & x(m) <= 0.75
      F(m) = 3 - 4 * x(m); end; end
% at t = 0, the solution is:
  tplot(1) = 0; u = zeros(M,N+1); u(1:M,1) = F(1:M);
% at t = Δt, the solution is given by Equation 7.5.16
  tplot(2) = dt;
  for m = 2:M-1
```

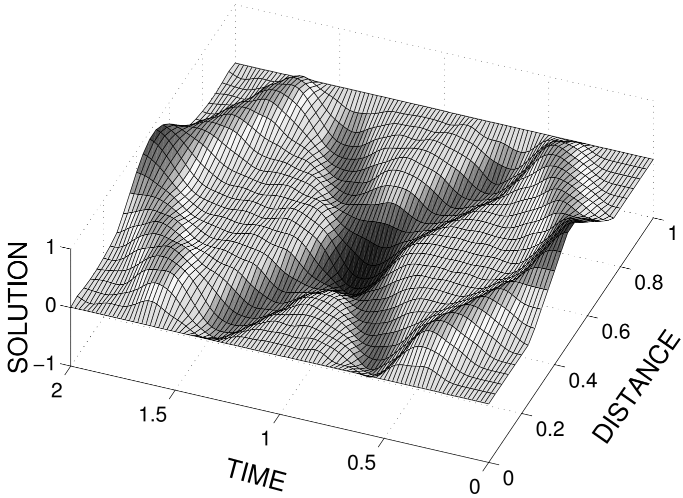


Figure 7.5.3: The numerical solution $u(x,t)/h$ of the wave equation with $c\Delta t/\Delta x = \frac{1}{2}$ using Equation 7.5.11 at various positions $x' = x/L$ and times $t' = ct/L$. The exact solution is plotted in Figure 7.3.2.

```
    u(m,2) = 0.5*coeffsq*(F(m+1)+F(m-1)) + (1-coeffsq)*F(m)+dt*G(m);
  end
% in general, the solution is given by Equation 7.5.11
  for n = 2:N
    tplot(n+1) = dt * n;
  for m = 2:M-1
    u(m,n+1) = 2*u(m,n)-u(m,n-1) + coeffsq*(u(m+1,n)-2*u(m,n)+u(m-1,n));
end; end
X = x' * ones(1,length(tplot)); T = ones(M,1) * tplot;
surf(X,T,u)
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

Overall, the numerical solution shown in Figure 7.5.3 approximates the exact or analytic solution well. However, we note small-scale noise in the numerical solution at later times. Why does this occur? Recall that the exact solution could be written as an infinite sum of sines in the x dimension. Each successive harmonic adds a contribution from waves of shorter and shorter wavelength. In the case of the numerical solution, the longer-wavelength harmonics are well represented by the numerical scheme because there are many grid points available to resolve a given wavelength. As the wavelengths become shorter, the higher harmonics are poorly resolved by the numerical scheme, move at incorrect phase speeds, and their misplacement (dispersion) creates the small-scale noise that you observe rather than giving the sharp angular features of the exact solution. The only method for avoiding this problem is to devise schemes that minimize dispersion. □

• **Example 7.5.2**

Let us redo Example 7.5.1 except that we introduce the boundary condition that $u_x(L,t) = 0$. This corresponds to a string where we fix the left end and allow the right end to freely move up and down. This requires a new difference condition along the right

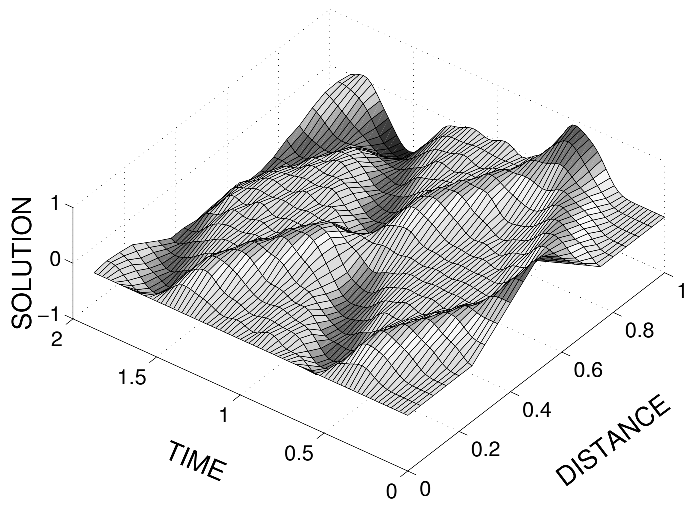


Figure 7.5.4: The numerical solution $u(x,t)/h$ of the wave equation when the right end moves freely with $c\Delta t/\Delta x = \frac{1}{2}$ using Equation 7.5.11 and Equation 7.5.30 at various positions $x' = x/L$ and times $t' = ct/L$.

boundary. If we employ centered differencing,

$$\frac{u_{L+1}^n - u_{L-1}^n}{2\Delta x} = 0, \tag{7.5.28}$$

and

$$u_L^{n+1} = 2u_L^n - u_L^{n-1} + \left(\frac{c\Delta t}{\Delta x}\right)^2 \left(u_{L+1}^n - 2u_L^n + u_{L-1}^n\right). \tag{7.5.29}$$

Eliminating u_{L+1}^n between Equation 7.5.28 and Equation 7.5.29,

$$u_L^{n+1} = 2u_L^n - u_L^{n-1} + \left(\frac{c\Delta t}{\Delta x}\right)^2 \left(2u_{L-1}^n - 2u_L^n\right). \tag{7.5.30}$$

For the special case of $n = 1$, Equation 7.5.30 becomes

$$u_L^1 = f(x_L) + \left(\frac{c\Delta t}{\Delta x}\right)^2 [f(x_{L-1}) - f(x_L)] + \Delta t f(x_L). \tag{7.5.31}$$

The MATLAB code used to numerically solve the wave equation with a Neumann boundary condition is very similar to the one used in the previous example that we must add the line

```
u(M,2) = coeffsq * F(M-1) + (1-coeffsq) * F(M) + dt*G(M);
```

after

```
for m = 2:M-1
  u(m,2) = 0.5*coeffsq*(F(m+1)+F(m-1)) + (1-coeffsq)*F(m)+dt*G(m);
end
```

and

```
  u(M,n+1) = 2*u(M,n)-u(M,n-1) + 2*coeffsq*(u(M-1,n)-u(M,n));
```

after

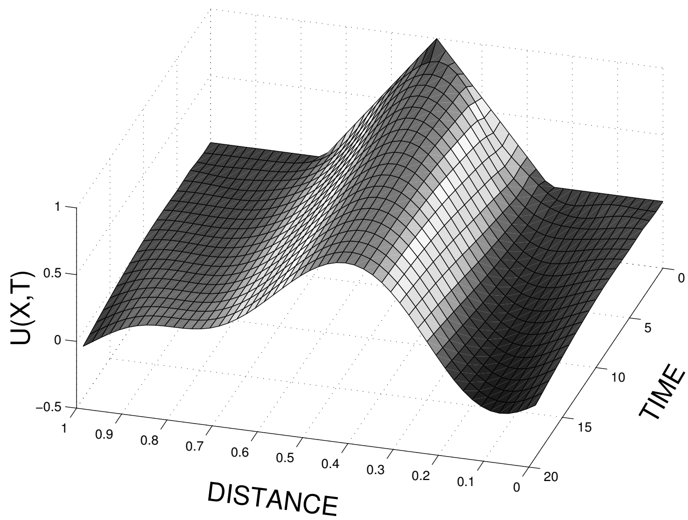


Figure 7.5.5: The numerical solution $u(x,t)$ of the first-order hyperbolic partial differential equation $u_t + u_x = 0$ using the Lax-Wendroff formula as observed at $t = 0, 1, 2, \ldots, 20$. The initial conditions are given by Equation 7.3.25 with $h = 1$, $\Delta t/\Delta x = \frac{2}{3}$, and $\Delta x = 0.02$.

```
for m = 2:M-1
  u(m,n+1) = 2*u(m,n)-u(m,n-1) + coeffsq*(u(m+1,n)-2*u(m,n)+u(m-1,n));
end
```

Figure 7.5.4 shows the results. The numerical solution agrees well with the exact solution

$$u(x,t) = \frac{32h}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \sin\left[\frac{(2n-1)\pi x}{2L}\right] \cos\left[\frac{(2n-1)\pi ct}{2L}\right] \times \left\{2\sin\left[\frac{(2n-1)\pi}{4}\right] - \sin\left[\frac{3(2n-1)\pi}{8}\right] - \sin\left[\frac{(2n-1)\pi}{8}\right]\right\}. \quad \textbf{(7.5.32)}$$

The results are also consistent with those presented in Example 7.5.1, especially with regard to small-scale noise due to dispersion.

Project: Numerical Solution of First-Order Hyperbolic Equations

The equation $u_t + u_x = 0$ is the simplest possible hyperbolic partial differential equation. Indeed, the classic wave equation consists of a system of these equations: $u_t + cv_x = 0$, and $v_t + cu_x = 0$. In this project you will examine several numerical schemes for solving such a partial differential equation using MATLAB.

Step 1: One of the simplest numerical schemes is the forward-in-time, centered-in-space of

$$\frac{u_m^{n+1} - u_m^n}{\Delta t} + \frac{u_{m+1}^n - u_{m-1}^n}{2\Delta x} = 0.$$

Use von Neumann's stability analysis to show that this scheme is *always* unstable.

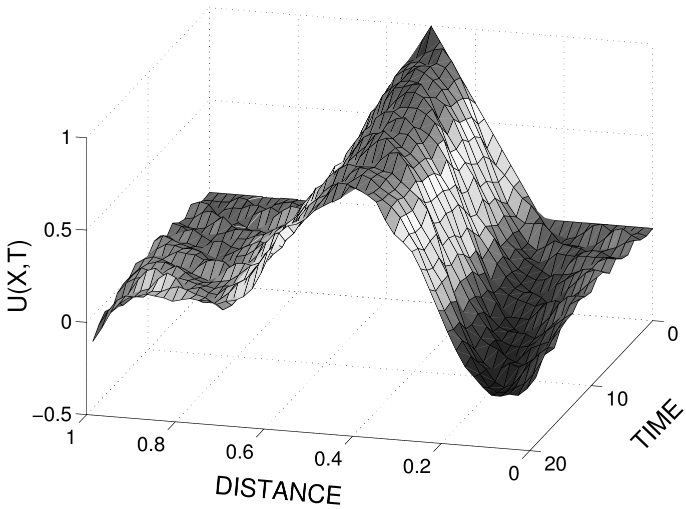


Figure 7.5.6: Same as Figure 7.5.5 except that the centered-in-time, centered-in-space scheme was used.

Step 2: The most widely used method for numerically integrating first-order hyperbolic equations is the *Lax-Wendroff* method:[18]

$$u_m^{n+1} = u_m^n - \frac{\Delta t}{2\Delta x}\left(u_{m+1}^n - u_{m-1}^n\right) + \frac{(\Delta t)^2}{2(\Delta x)^2}\left(u_{m+1}^n - 2u_m^n + u_{m-1}^n\right).$$

This method introduces errors of $O[(\Delta t)^2]$ and $O[(\Delta x)^2]$. Show that this scheme is stable if it satisfies the CFL criteria of $\Delta t/\Delta x \le 1$.

Using the initial condition given by Equation 7.3.25, write a MATLAB code that uses this scheme to numerically integrate $u_t + u_x = 0$. Plot the results for various $\Delta t/\Delta x$ over the interval $0 \le x \le 1$ given the *periodic* boundary conditions of $u(0,t) = u(1,t)$ for the temporal interval $0 \le t \le 20$. See Figure 7.5.5. Discuss the strengths and weaknesses of the scheme with respect to dissipation or damping of the numerical solution and preserving the phase of the solution. Most numerical methods books discuss this.[19]

Step 3: Another simple scheme is the centered-in-time, centered-in-space of

$$\frac{u_m^{n+1} - u_m^{n-1}}{2\Delta t} + \frac{u_{m+1}^n - u_{m-1}^n}{2\Delta x} = 0.$$

This method introduces errors of $O[(\Delta t)^2]$ and $O[(\Delta x)^2]$.

Repeat the analysis from Step 1 for this scheme. One of the difficulties is taking the first time step. Use the scheme in Step 1 to take this first time step. See Figure 7.5.6.

Further Reading

King, G. C., 2009: *Vibrations and Waves.* Wiley, 228 pp. This book emphasises the physical principles, rather than the mathematics.

[18] Lax, P. D., and B. Wendroff, 1960: Systems of conservative laws. *Comm. Pure Appl. Math.*, **13**, 217–237.

[19] For example, Lapidus, L., and G. F. Pinder, 1982: *Numerical Solution of Partial Differential Equations in Science and Engineering*. John Wiley & Sons, 677 pp.

Koshlyakov, N. S., M. M. Smirnov, and E. B. Gliner, 1964: *Differential Equations of Mathematical Physics.* North-Holland Publishing, 701 pp. See Part I. Detailed presentations of solution techniques.

Morse, P. M., and H. Feshback, 1953: *Methods of Theoretical Physics.* McGraw -Hill Book Co., 997 pp. Chapter 11 is devoted to solving the wave equation.

Chapter 8
The Heat Equation

In this chapter we deal with the linear parabolic differential equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} \tag{8.0.1}$$

in the two independent variables x and t. This equation, known as the one-dimensional heat equation, serves as the prototype for a wider class of *parabolic equations*

$$a(x,t)\frac{\partial^2 u}{\partial x^2} + b(x,t)\frac{\partial^2 u}{\partial x \partial t} + c(x,t)\frac{\partial^2 u}{\partial t^2} = f\left(x, t, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial t}\right), \tag{8.0.2}$$

where $b^2 = 4ac$. It arises in the study of heat conduction in solids as well as in a variety of diffusive phenomena. The heat equation is similar to the wave equation in that it is also an equation of evolution. However, the heat equation is not "conservative" because if we reverse the sign of t, we obtain a different solution. This reflects the presence of entropy, which must always increase during heat conduction.

8.1 DERIVATION OF THE HEAT EQUATION

To derive the heat equation, consider a heat-conducting homogeneous rod, extending from $x = 0$ to $x = L$ along the x-axis (see Figure 8.1.1). The rod has uniform cross section A and constant density ρ, is insulated laterally so that heat flows only in the x-direction, and is sufficiently thin so that the temperature at all points on a cross section is constant. Let $u(x,t)$ denote the temperature of the cross section at the point x at any instant of time t, and let c denote the specific heat of the rod (the amount of heat required to raise the

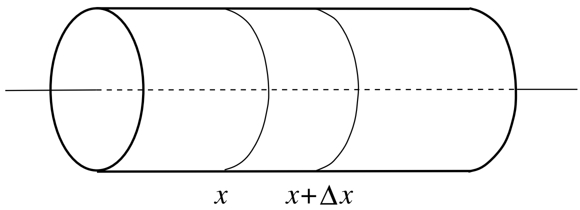


Figure 8.1.1: Heat conduction in a thin bar.

temperature of a unit mass of the rod by a degree). In the segment of the rod between the cross section at x and the cross section at $x + \Delta x$, the amount of heat is

$$Q(t) = \int_x^{x+\Delta x} c\rho A u(s,t)\, ds. \qquad \textbf{(8.1.1)}$$

On the other hand, the rate at which heat flows into the segment across the cross section at x is proportional to the cross section and the gradient of the temperature at the cross section (Fourier's law of heat conduction):

$$-\kappa A \frac{\partial u(x,t)}{\partial x}, \qquad \textbf{(8.1.2)}$$

where κ denotes the thermal conductivity of the rod. The sign in Equation 8.1.2 indicates that heat flows in the direction of decreasing temperature. Similarly, the rate at which heat flows out of the segment through the cross section at $x + \Delta x$ equals

$$-\kappa A \frac{\partial u(x+\Delta x,t)}{\partial x}. \qquad \textbf{(8.1.3)}$$

The difference between the amount of heat that flows in through the cross section at x and the amount of heat that flows out through the cross section at $x+\Delta x$ must equal the change in the heat content of the segment $x \le s \le x + \Delta x$. Hence, by subtracting Equation 8.1.3 from Equation 8.1.2 and equating the result to the time derivative of Equation 8.1.1,

$$\frac{\partial Q}{\partial t} = \int_x^{x+\Delta x} c\rho A \frac{\partial u(s,t)}{\partial t}\, ds = \kappa A \left[\frac{\partial u(x+\Delta x,t)}{\partial x} - \frac{\partial u(x,t)}{\partial x} \right]. \qquad \textbf{(8.1.4)}$$

Assuming that the integrand in Equation 8.1.4 is a continuous function of s, then by the mean value theorem for integrals,

$$\int_x^{x+\Delta x} \frac{\partial u(s,t)}{\partial t}\, ds = \frac{\partial u(\xi,t)}{\partial t} \Delta x, \qquad x < \xi < x + \Delta x, \qquad \textbf{(8.1.5)}$$

so that Equation 8.1.4 becomes

$$c\rho \Delta x \frac{\partial u(\xi,t)}{\partial t} = \kappa \left[\frac{\partial u(x+\Delta x,t)}{\partial x} - \frac{\partial u(x,t)}{\partial x} \right]. \qquad \textbf{(8.1.6)}$$

Dividing both sides of Equation 8.1.6 by $c\rho\Delta x$ and taking the limit as $\Delta x \to 0$,

$$\frac{\partial u(x,t)}{\partial t} = a^2 \frac{\partial^2 u(x,t)}{\partial x^2} \qquad \textbf{(8.1.7)}$$

with $a^2 = \kappa/(c\rho)$. Equation 8.1.7 is called the one-dimensional *heat equation.* The constant a^2 is called the *diffusivity* within the solid.

If an external source supplies heat to the rod at a rate $f(x,t)$ per unit volume per unit time, we must add the term $\int_x^{x+\Delta x} f(s,t)\,ds$ to the time derivative term of Equation 8.1.4. Thus, in the limit $\Delta x \to 0$,

$$\frac{\partial u(x,t)}{\partial t} - a^2 \frac{\partial^2 u(x,t)}{\partial x^2} = F(x,t), \tag{8.1.8}$$

where $F(x,t) = f(x,t)/(c\rho)$ is the source density. This equation is called the *nonhomogeneous heat equation.*

8.2 INITIAL AND BOUNDARY CONDITIONS

In the case of heat conduction in a thin rod, the temperature function $u(x,t)$ must satisfy not only the heat equation, Equation 8.1.7, but also how the two ends of the rod exchange heat energy with the surrounding medium. If (1) there is no heat source, (2) the function $f(x)$, $0 < x < L$ describes the temperature in the rod at $t = 0$, and (3) we maintain both ends at zero temperature for all time, then the partial differential equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{8.2.1}$$

describes the temperature distribution $u(x,t)$ in the rod at any later time $0 < t$ subject to the conditions

$$u(x,0) = f(x), \qquad 0 < x < L, \tag{8.2.2}$$

and

$$u(0,t) = u(L,t) = 0, \qquad 0 < t. \tag{8.2.3}$$

Equations 8.2.1 and 8.2.3 describe the *initial-boundary-value problem* for this particular heat conduction problem; Equation 8.2.3 is the boundary condition while Equation 8.2.2 gives the initial condition. Note that in the case of the heat equation, the problem only demands the initial value of $u(x,t)$ and not $u_t(x,0)$, as with the wave equation.

Historically most linear boundary conditions have been classified in one of three ways. The condition, Equation 8.2.3, is an example of a *Dirichlet problem*[1] or *condition of the first kind*. This type of boundary condition gives the value of the solution (which is not necessarily equal to zero) along a boundary.

The next simplest condition involves derivatives. If we insulate both ends of the rod so that no heat flows from the ends, then according to Equation 7.1.2 the boundary condition assumes the form

$$\frac{\partial u(0,t)}{\partial x} = \frac{\partial u(L,t)}{\partial x} = 0, \qquad 0 < t. \tag{8.2.4}$$

This is an example of a *Neumann problem*[2] or *condition of the second kind*. This type of boundary condition specifies the value of the normal derivative (which may not be equal to zero) of the solution along the boundary.

---

[1] Dirichlet, P. G. L., 1850: Über einen neuen Ausdruck zur Bestimmung der Dichtigkeit einer unendlich dünnen Kugelschale, wenn der Werth des Potentials derselben in jedem Punkte ihrer Oberfläche gegeben ist. *Abh. Königlich. Preuss. Akad. Wiss.*, 99–116.

[2] Neumann, C. G., 1877: *Untersuchungen über das Logarithmische und Newton'sche Potential.*

Finally, if there is radiation of heat from the ends of the rod into the surrounding medium, we shall show that the boundary condition is of the form

$$\frac{\partial u(0,t)}{\partial x} - hu(0,t) = \text{a constant}, \tag{8.2.5}$$

and

$$\frac{\partial u(L,t)}{\partial x} + hu(L,t) = \text{another constant} \tag{8.2.6}$$

for $0 < t$, where h is a positive constant. This is an example of a *condition of the third kind* or *Robin problem*[3] and is a linear combination of Dirichlet and Neumann conditions.

8.3 SEPARATION OF VARIABLES

As with the wave equation, the most popular and widely used technique for solving the heat equation is separation of variables. Its success depends on our ability to express the solution $u(x,t)$ as the product $X(x)T(t)$. If we cannot achieve this separation, then the technique must be abandoned for others. In the following examples we show how to apply this technique.

• Example 8.3.1

Let us find the solution to the homogeneous heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{8.3.1}$$

which satisfies the initial condition

$$u(x,0) = f(x), \quad 0 < x < L, \tag{8.3.2}$$

and the boundary conditions

$$u(0,t) = u(L,t) = 0, \quad 0 < t. \tag{8.3.3}$$

This system of equations models heat conduction in a thin metallic bar where both ends are held at the constant temperature of zero and the bar initially has the temperature $f(x)$.

We shall solve this problem by the method of separation of variables. Accordingly, we seek particular solutions of Equation 8.3.1 of the form

$$u(x,t) = X(x)T(t), \tag{8.3.4}$$

which satisfy the boundary conditions, Equation 8.3.3. Because

$$\frac{\partial u}{\partial t} = X(x)T'(t), \tag{8.3.5}$$

and

$$\frac{\partial^2 u}{\partial x^2} = X''(x)T(t), \tag{8.3.6}$$

[3] Robin, G., 1886: Sur la distribution de l'électricité à la surface des conducteurs fermés et des conducteurs ouverts. *Ann. Sci. l'Ecole Norm. Sup., Ser. 3*, **3**, S1–S58.

Equation 8.3.1 becomes

$$T'(t)X(x) = a^2X''(x)T(t). \tag{8.3.7}$$

Dividing both sides of Equation 8.3.7 by $a^2X(x)T(t)$ gives

$$\frac{T'}{a^2T} = \frac{X''}{X} = -\lambda, \tag{8.3.8}$$

where $-\lambda$ is the separation constant. Equation 8.3.8 immediately yields two ordinary differential equations:

$$X'' + \lambda X = 0, \tag{8.3.9}$$

and

$$T' + a^2\lambda T = 0 \tag{8.3.10}$$

for the functions $X(x)$ and $T(t)$, respectively.

We now rewrite the boundary conditions in terms of $X(x)$ by noting that the boundary conditions are $u(0,t) = X(0)T(t) = 0$, and $u(L,t) = X(L)T(t) = 0$ for $0 < t$. If we were to choose $T(t) = 0$, then we would have a trivial solution for $u(x,t)$. Consequently, $X(0) = X(L) = 0$.

We now solve Equation 8.3.9. There are three possible cases: $\lambda = -m^2$, $\lambda = 0$, and $\lambda = k^2$. If $\lambda = -m^2 < 0$, then we must solve the boundary-value problem

$$X'' - m^2X = 0, \qquad X(0) = X(L) = 0. \tag{8.3.11}$$

The general solution to Equation 8.3.11 is

$$X(x) = A\cosh(mx) + B\sinh(mx). \tag{8.3.12}$$

Because $X(0) = 0$, it follows that $A = 0$. The condition $X(L) = 0$ yields $B\sinh(mL) = 0$. Since $\sinh(mL) \neq 0$, $B = 0$, and we have a trivial solution for $\lambda < 0$.

If $\lambda = 0$, the corresponding boundary-value problem is

$$X''(x) = 0, \qquad X(0) = X(L) = 0. \tag{8.3.13}$$

The general solution is

$$X(x) = C + Dx. \tag{8.3.14}$$

From $X(0) = 0$, we have that $C = 0$. From $X(L) = 0$, $DL = 0$, or $D = 0$. Again, we obtain a trivial solution.

Finally, we assume that $\lambda = k^2 > 0$. The corresponding boundary-value problem is

$$X'' + k^2X = 0, \qquad X(0) = X(L) = 0. \tag{8.3.15}$$

The general solution to Equation 8.3.15 is

$$X(x) = E\cos(kx) + F\sin(kx). \tag{8.3.16}$$

Because $X(0) = 0$, it follows that $E = 0$; from $X(L) = 0$, we obtain $F\sin(kL) = 0$. For a nontrivial solution, $F \neq 0$ and $\sin(kL) = 0$. This implies that $k_nL = n\pi$, where $n = 1, 2, 3, \ldots$. In summary, the x-dependence of the solution is

$$X_n(x) = F_n \sin\left(\frac{n\pi x}{L}\right), \tag{8.3.17}$$

where $\lambda_n = n^2\pi^2/L^2$.

Turning to the time dependence, we use $\lambda_n = n^2\pi^2/L^2$ in Equation 8.3.10

$$T_n' + \frac{a^2n^2\pi^2}{L^2}T_n = 0. \tag{8.3.18}$$

The corresponding general solution is

$$T_n(t) = G_n \exp\left(-\frac{a^2n^2\pi^2}{L^2}t\right). \tag{8.3.19}$$

Thus, the functions

$$u_n(x,t) = B_n \sin\left(\frac{n\pi x}{L}\right)\exp\left(-\frac{a^2n^2\pi^2}{L^2}t\right), n = 1, 2, 3, \ldots, \tag{8.3.20}$$

where $B_n = F_nG_n$, are particular solutions of Equation 8.3.1 and satisfy the homogeneous boundary conditions, Equation 8.3.3.

As we noted in the case of the wave equation, we can solve the x-dependence equation as a regular Sturm-Liouville problem. After finding the eigenvalue λ_n and eigenfunction, we solve for $T_n(t)$. The product solution $u_n(x,t)$ equals the product of the eigenfunction and $T_n(t)$.

Having found particular solutions to our problem, the most general solution equals a linear sum of these particular solutions:

$$u(x,t) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right)\exp\left(-\frac{a^2n^2\pi^2}{L^2}t\right). \tag{8.3.21}$$

The coefficient B_n is chosen so that Equation 8.3.21 yields the initial condition, Equation 8.3.2, if $t = 0$. Thus, setting $t = 0$ in Equation 8.3.21, we see from Equation 8.3.2 that the coefficients B_n must satisfy the relationship

$$f(x) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right), \qquad 0 < x < L. \tag{8.3.22}$$

This is precisely a Fourier half-range sine series for $f(x)$ on the interval $(0, L)$. Therefore, the formula

$$B_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{n\pi x}{L}\right)\,dx, \qquad n = 1, 2, 3, \ldots \tag{8.3.23}$$

gives the coefficients B_n. For example, if $L = \pi$ and $u(x,0) = x(\pi - x)$, then

$$B_n = \frac{2}{\pi}\int_0^\pi x(\pi - x)\sin(nx)\,dx = 2\int_0^\pi x\sin(nx)\,dx - \frac{2}{\pi}\int_0^\pi x^2\sin(nx)\,dx = 4\frac{1-(-1)^n}{n^3\pi}. \tag{8.3.24}$$

Hence,

$$u(x,t) = \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(2n-1)x]}{(2n-1)^3}e^{-(2n-1)^2a^2t}. \tag{8.3.25}$$

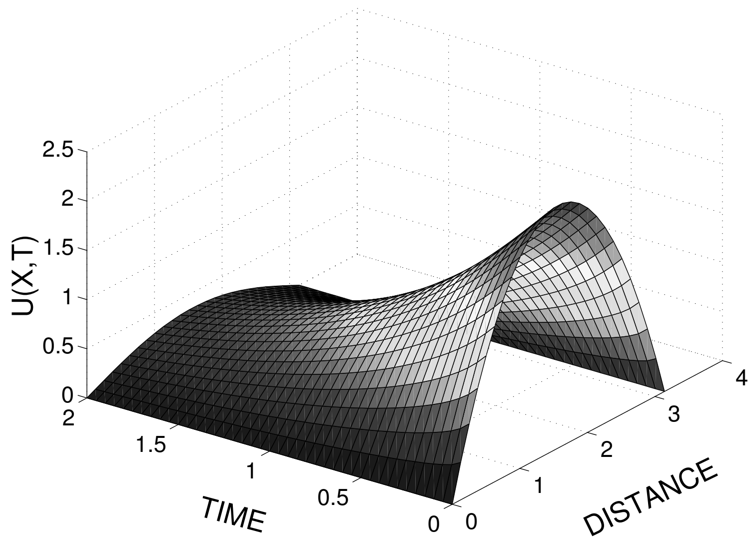


Figure 8.3.1: The temperature $u(x,t)$ within a thin bar as a function of position x and time a^2t when we maintain both ends at zero and the initial temperature equals $x(\pi - x)$.

Figure 8.3.1 illustrates Equation 8.3.25 for various times. It was created using the MATLAB script

```
clear
M = 20; dx = pi/25; dt = 0.05;
% compute grid and initialize solution
X = [0:dx:pi]; T = [0:dt:2];
u = zeros(length(T),length(X));
XX = repmat(X,[length(T) 1]); TT = repmat(T',[1 length(X)]);
% compute solution from Equation 8.3.25
for m = 1:M
  temp1 = 2*m-1; coeff = 8 / (pi * temp1 * temp1 * temp1);
  u = u + coeff * sin(temp1*XX) .* exp(-temp1 * temp1 * TT);
end
surf(XX,TT,u)
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('U(X,T)','Fontsize',20)
```

Note that both ends of the bar satisfy the boundary conditions, namely that the temperature equals zero. As time increases, heat flows out from the center of the bar to both ends where it is removed. This process is reflected in the collapse of the original parabolic shape of the temperature profile toward zero as time increases. □

• **Example 8.3.2**

As a second example, let us solve the heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{8.3.26}$$

which satisfies the initial condition

$$u(x,0) = x, \quad 0 < x < L, \tag{8.3.27}$$

and the boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = u(L,t) = 0, \quad 0 < t. \tag{8.3.28}$$

The condition $u_x(0,t) = 0$ expresses mathematically the constraint that no heat flows through the left boundary (insulated end condition).

Once again, we employ separation of variables; as in the previous example, the positive and zero separation constants yield trivial solutions. For a negative separation constant, however,

$$X'' + k^2 X = 0, \tag{8.3.29}$$

with

$$X'(0) = X(L) = 0, \tag{8.3.30}$$

because $u_x(0,t) = X'(0)T(t) = 0$, and $u(L,t) = X(L)T(t) = 0$. This regular Sturm-Liouville problem has the solution

$$X_n(x) = \cos\left[\frac{(2n-1)\pi x}{2L}\right], \qquad n = 1, 2, 3, \ldots. \tag{8.3.31}$$

The temporal solution then becomes

$$T_n(t) = B_n \exp\left[-\frac{a^2(2n-1)^2\pi^2 t}{4L^2}\right]. \tag{8.3.32}$$

Consequently, a linear superposition of the particular solutions gives the total solution, which equals

$$u(x,t) = \sum_{n=1}^{\infty} B_n \cos\left[\frac{(2n-1)\pi x}{2L}\right] \exp\left[-\frac{a^2(2n-1)^2\pi^2}{4L^2} t\right]. \tag{8.3.33}$$

Our final task remains to find the coefficients B_n. Evaluating Equation 8.3.33 at $t = 0$,

$$u(x,0) = x = \sum_{n=1}^{\infty} B_n \cos\left[\frac{(2n-1)\pi x}{2L}\right], \qquad 0 < x < L. \tag{8.3.34}$$

Equation 8.3.34 is not a half-range cosine expansion; it is an expansion in the orthogonal functions $\cos[(2n-1)\pi x/(2L)]$ corresponding to the regular Sturm-Liouville problem, Equation 8.3.29 and Equation 8.3.30. Consequently, B_n is given by Equation 6.3.4 with $r(x) = 1$ as

$$B_n = \frac{\int_0^L x \cos[(2n-1)\pi x/(2L)]\, dx}{\int_0^L \cos^2[(2n-1)\pi x/(2L)]\, dx} \tag{8.3.35}$$

$$= \frac{\frac{4L^2}{(2n-1)^2\pi^2} \cos\left[\frac{(2n-1)\pi x}{2L}\right]\Big|_0^L + \frac{2Lx}{(2n-1)\pi} \sin\left[\frac{(2n-1)\pi x}{2L}\right]\Big|_0^L}{\frac{x}{2}\Big|_0^L + \frac{L}{2(2n-1)\pi} \sin\left[\frac{(2n-1)\pi x}{L}\right]\Big|_0^L} \tag{8.3.36}$$

$$= \frac{8L}{(2n-1)^2\pi^2}\left\{\cos\left[\frac{(2n-1)\pi}{2}\right] - 1\right\} + \frac{4L}{(2n-1)\pi} \sin\left[\frac{(2n-1)\pi}{2}\right] \tag{8.3.37}$$

$$= -\frac{8L}{(2n-1)^2\pi^2} - \frac{4L(-1)^n}{(2n-1)\pi}, \tag{8.3.38}$$

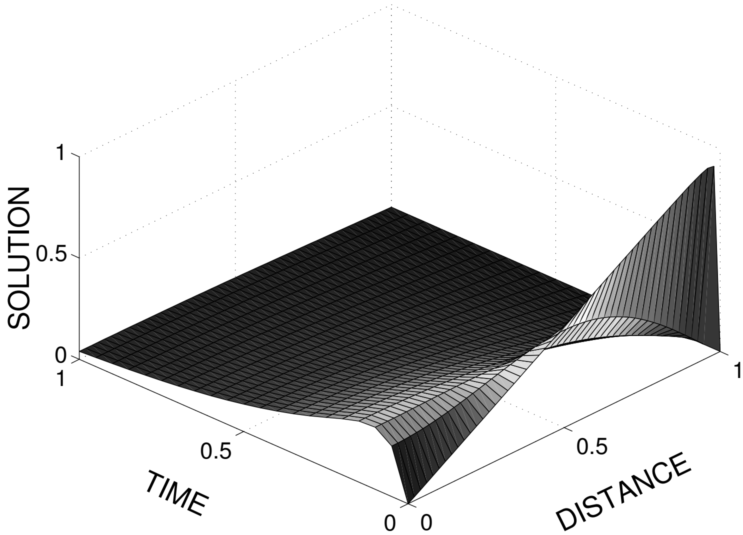


Figure 8.3.2: The temperature $u(x,t)/L$ within a thin bar as a function of position x/L and time a^2t/L^2 when we insulate the left end and hold the right end at the temperature of zero. The initial temperature equals x.

as $\cos[(2n-1)\pi/2] = 0$, and $\sin[(2n-1)\pi/2] = (-1)^{n+1}$. Consequently, the complete solution is

$$u(x,t) = -\frac{4L}{\pi}\sum_{n=1}^{\infty}\left[\frac{2}{(2n-1)^2\pi} + \frac{(-1)^n}{2n-1}\right]\cos\left[\frac{(2n-1)\pi x}{2L}\right]\exp\left[-\frac{(2n-1)^2\pi^2a^2t}{4L^2}\right]. \tag{8.3.39}$$

Figure 8.3.2 illustrates the evolution of the temperature field with time. It was generated using the MATLAB script

```
clear
M = 200; dx = 0.02; dt = 0.05;
% compute Fourier coefficients
sign = -1;
for m = 1:M
  temp1 = 2*m-1;
  a(m) = 2/(pi*temp1*temp1) + sign/temp1;
  sign = - sign;
end
% compute grid and initialize solution
X = [0:dx:1]; T = [0:dt:1];
u = zeros(length(T),length(X));
XX = repmat(X,[length(T) 1]);
TT = repmat(T',[1 length(X)]);
% compute solution from Equation 8.3.39
for m = 1:M
  temp1 = (2*m-1)*pi/2;
  u = u + a(m) * cos(temp1*XX) .* exp(-temp1 * temp1 * TT);
end
```

```
u = - (4/pi) * u;
surf(XX,TT,u); axis([0 1 0 1 0 1]);
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

Initially, heat near the center of the bar flows toward the cooler, insulated end, resulting in an increase of temperature there. On the right side, heat flows out of the bar because the temperature is maintained at zero at $x = L$. Eventually the heat that has accumulated at the left end flows rightward because of the continual heat loss on the right end. In the limit of $t \to \infty$, all of the heat has left the bar. □

• Example 8.3.3

A slight variation on Example 8.3.1 is

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{8.3.40}$$

where

$$u(x,0) = u(0,t) = 0, \quad \text{and} \quad u(L,t) = \theta. \tag{8.3.41}$$

We begin by blindly employing the technique of separation of variables. Once again, we obtain the ordinary differential equation, Equation 8.3.9 and Equation 8.3.10. The initial and boundary conditions become, however,

$$X(0) = T(0) = 0, \tag{8.3.42}$$

and

$$X(L)T(t) = \theta. \tag{8.3.43}$$

Although Equation 8.3.42 is acceptable, Equation 8.3.43 gives us an impossible condition because $T(t)$ cannot be constant. If it were, it would have to equal to zero by Equation 8.3.42.

To find a way around this difficulty, suppose that we want the solution to our problem at a time long after $t = 0$. From experience we know that heat conduction with time-independent boundary conditions eventually results in an evolution from the initial condition to some time-independent (steady-state) equilibrium. If we denote this steady-state solution by $w(x)$, it must satisfy the heat equation

$$a^2 w''(x) = 0, \tag{8.3.44}$$

and the boundary conditions

$$w(0) = 0, \quad \text{and} \quad w(L) = \theta. \tag{8.3.45}$$

We can integrate Equation 8.3.44 immediately to give

$$w(x) = A + Bx, \tag{8.3.46}$$

and the boundary condition, Equation 8.3.45, results in

$$w(x) = \frac{\theta x}{L}. \tag{8.3.47}$$

Clearly Equation 8.3.47 cannot hope to satisfy the initial conditions; that was never expected of it. However, if we add a time-varying (transient) solution $v(x,t)$ to $w(x)$ so that

$$u(x,t) = w(x) + v(x,t), \tag{8.3.48}$$

we could satisfy the initial condition if

$$v(x,0) = u(x,0) - w(x), \tag{8.3.49}$$

and $v(x,t)$ tends to zero as $t \to \infty$. Furthermore, because $w''(x) = w(0) = 0$, and $w(L) = \theta$,

$$\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 v}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{8.3.50}$$

with the boundary conditions

$$v(0,t) = v(L,t) = 0, \quad 0 < t. \tag{8.3.51}$$

We can solve Equation 8.3.49, Equation 8.3.50, and Equation 8.3.51 by separation of variables; we did it in Example 8.3.1. However, in place of $f(x)$ we now have $u(x,0) - w(x)$, or $-w(x)$ because $u(x,0) = 0$. Therefore, the solution $v(x,t)$ is

$$v(x,t) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right) \exp\left(-\frac{a^2 n^2 \pi^2}{L^2} t\right) \tag{8.3.52}$$

with

$$B_n = \frac{2}{L}\int_0^L -w(x) \sin\left(\frac{n\pi x}{L}\right) dx = \frac{2}{L}\int_0^L -\frac{\theta x}{L} \sin\left(\frac{n\pi x}{L}\right) dx \tag{8.3.53}$$

$$= -\frac{2\theta}{L^2}\left[\frac{L^2}{n^2\pi^2}\sin\left(\frac{n\pi x}{L}\right) - \frac{xL}{n\pi}\cos\left(\frac{n\pi x}{L}\right)\right]_0^L = (-1)^n \frac{2\theta}{n\pi}. \tag{8.3.54}$$

Thus, the entire solution is

$$u(x,t) = \frac{\theta x}{L} + \frac{2\theta}{\pi}\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{n\pi x}{L}\right) \exp\left(-\frac{a^2 n^2 \pi^2}{L^2} t\right). \tag{8.3.55}$$

The quantity $a^2 t/L^2$ is the *Fourier number*.

Figure 8.3.3 illustrates our solution and was created with the MATLAB script

```
clear
M = 1000; dx = 0.01; dt = 0.01;
% compute grid and initialize solution
X = [0:dx:1]; T = [0:dt:0.2];
XX = repmat(X,[length(T) 1]); TT = repmat(T',[1 length(X)]);
u = XX;
% compute solution from Equation 8.3.55
sign = -2/pi;
for m = 1:M
  coeff = sign/m;
```

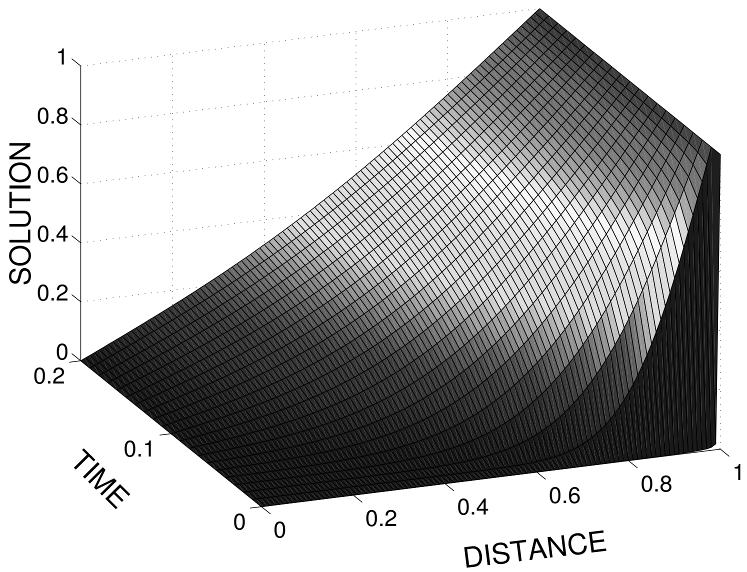


Figure 8.3.3: The temperature $u(x,t)/\theta$ within a thin bar as a function of position x/L and time a^2t/L^2 with the left end held at a temperature of zero and right end held at a temperature θ while the initial temperature of the bar is zero.

```
  u = u + coeff * sin((m*pi)*XX) .* exp(-(m*m*pi*pi) * TT);
  sign = -sign;
end
surf(XX,TT,u); axis([0 1 0 0.2 0 1]);
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

Clearly it satisfies the boundary conditions. Initially, heat flows rapidly from right to left. As time increases, the rate of heat transfer decreases until the final equilibrium (steady-state) is established and no more heat flows. □

• Example 8.3.4

Let us find the solution to the heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{8.3.56}$$

subject to the Neumann boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = \frac{\partial u(L,t)}{\partial x} = 0, \qquad 0 < t, \tag{8.3.57}$$

and the initial condition that

$$u(x,0) = x, \qquad 0 < x < L. \tag{8.3.58}$$

We have now insulated *both* ends of the bar.

Assuming that $u(x,t) = X(x)T(t)$,

$$\frac{T'}{a^2 T} = \frac{X''}{X} = -k^2, \tag{8.3.59}$$

where we have presently assumed that the separation constant is negative. The Neumann conditions give $u_x(0,t) = X'(0)T(t) = 0$, and $u_x(L,t) = X'(L)T(t) = 0$ so that $X'(0) = X'(L) = 0$.

The Sturm-Liouville problem

$$X'' + k^2 X = 0, \tag{8.3.60}$$

and

$$X'(0) = X'(L) = 0 \tag{8.3.61}$$

gives the x-dependence. The eigenfunction solution is

$$X_n(x) = \cos\left(\frac{n\pi x}{L}\right), \tag{8.3.62}$$

where $k_n = n\pi/L$ and $n = 1, 2, 3, \ldots$.

The corresponding temporal part equals the solution of

$$T_n' + a^2 k_n^2 T_n = T_n' + \frac{a^2 n^2 \pi^2}{L^2} T_n = 0, \tag{8.3.63}$$

which is

$$T_n(t) = A_n \exp\left(-\frac{a^2 n^2 \pi^2}{L^2} t\right). \tag{8.3.64}$$

Thus, the product solution given by a negative separation constant is

$$u_n(x,t) = X_n(x)T_n(t) = A_n \cos\left(\frac{n\pi x}{L}\right) \exp\left(-\frac{a^2 n^2 \pi^2}{L^2} t\right). \tag{8.3.65}$$

Unlike our previous problems, there is a nontrivial solution for a separation constant that equals zero. In this instance, the x-dependence equals

$$X(x) = Ax + B. \tag{8.3.66}$$

The boundary conditions $X'(0) = X'(L) = 0$ force A to be zero but B is completely free. Consequently, the eigenfunction in this particular case is

$$X_0(x) = 1. \tag{8.3.67}$$

Because $T_0'(t) = 0$ in this case, the temporal part equals a constant that we shall take to be $A_0/2$. Therefore, the product solution corresponding to the zero separation constant is

$$u_0(x,t) = X_0(x)T_0(t) = A_0/2. \tag{8.3.68}$$

The most general solution to our problem equals the sum of all of the possible solutions:

$$u(x,t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi x}{L}\right) \exp\left(-\frac{a^2 n^2 \pi^2}{L^2} t\right). \tag{8.3.69}$$

Upon substituting $t = 0$ into Equation 8.3.69, we can determine A_n because

$$u(x,0) = x = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi x}{L}\right) \tag{8.3.70}$$

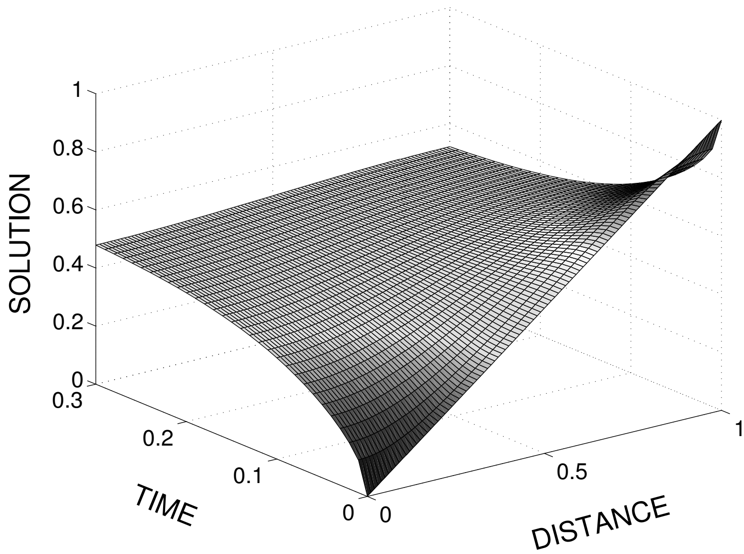


Figure 8.3.4: The temperature $u(x,t)/L$ within a thin bar as a function of position x/L and time a^2t/L^2 when we insulate both ends. The initial temperature of the bar is x.

is merely a half-range Fourier cosine expansion of the function x over the interval $(0, L)$. From Equation 5.1.22 and Equation 5.1.23,

$$A_0 = \frac{2}{L}\int_0^L x\,dx = L, \tag{8.3.71}$$

and

$$A_n = \frac{2}{L}\int_0^L x\cos\left(\frac{n\pi x}{L}\right)\,dx = \frac{2}{L}\left[\frac{L^2}{n^2\pi^2}\cos\left(\frac{n\pi x}{L}\right) + \frac{xL}{n\pi}\sin\left(\frac{n\pi x}{L}\right)\right]_0^L = \frac{2L}{n^2\pi^2}\left[(-1)^n - 1\right]. \tag{8.3.72}$$

The complete solution is

$$u(x,t) = \frac{L}{2} - \frac{4L}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^2}\cos\left[\frac{(2m-1)\pi x}{L}\right]\exp\left[-\frac{a^2(2m-1)^2\pi^2}{L^2}t\right], \tag{8.3.73}$$

because all of the even harmonics vanish and we may rewrite the odd harmonics using $n = 2m - 1$, where $m = 1, 2, 3, 4, \ldots$.

Figure 8.3.4 illustrates Equation 8.3.73 for various positions and times. It was generated using the MATLAB script

```
clear
M = 100; dx = 0.01; dt = 0.01;
% compute grid and initialize solution
X = [0:dx:1]; T = [0:dt:0.3];
u = zeros(length(T),length(X)); u = 0.5;
XX = repmat(X,[length(T) 1]); TT = repmat(T',[1 length(X)]);
% compute solution from Equation 8.3.73
for m = 1:M
  temp1 = (2*m-1) * pi;
```

```
  coeff = 4 / (temp1*temp1);
  u = u - coeff * cos(temp1*XX) .* exp(-temp1 * temp1 * TT);
end
surf(XX,TT,u); axis([0 1 0 0.3 0 1]);
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

The physical interpretation is quite simple. Since heat cannot flow in or out of the rod because of the insulation, it can only redistribute itself. Thus, heat flows from the warm right end to the cooler left end. Eventually the temperature achieves steady-state when the temperature is uniform throughout the bar. □

• **Example 8.3.5**

So far we have dealt with problems where the temperature or flux of heat has been specified at the ends of the rod. In many physical applications, one or both of the ends may radiate to free space at temperature u_0. According to Stefan's law, the amount of heat radiated from a given area dA in a given time interval dt is

$$\sigma(u^4 - u_0^4)\, dA\, dt, \tag{8.3.74}$$

where σ is called the Stefan-Boltzmann constant. On the other hand, the amount of heat that reaches the surface from the interior of the body, assuming that we are at the right end of the bar, equals

$$-\kappa \frac{\partial u}{\partial x}\, dA\, dt, \tag{8.3.75}$$

where κ is the thermal conductivity. Because these quantities must be equal,

$$-\kappa \frac{\partial u}{\partial x} = \sigma(u^4 - u_0^4) = \sigma(u - u_0)(u^3 + u^2 u_0 + u u_0^2 + u_0^3). \tag{8.3.76}$$

If u and u_0 are nearly equal, we may approximate the second bracketed term on the right side of Equation 8.3.76 as $4u_0^3$. We write this approximate form of Equation 8.3.76 as

$$-\frac{\partial u}{\partial x} = h(u - u_0), \tag{8.3.77}$$

where h, the *surface conductance* or the *coefficient of surface heat transfer*, equals $4\sigma u_0^3/\kappa$. Equation 8.3.77 is a "radiation" boundary condition. Sometimes someone will refer to it as "Newton's law" because Equation 8.3.77 is mathematically identical to Newton's law of cooling of a body by forced convection.

Let us now solve the problem of a rod that we initially heat to the uniform temperature of 100. We then allow it to cool by maintaining the temperature at zero at $x = 0$ and radiatively cooling to the surrounding air at the temperature of zero[4] at $x = L$. We may restate the problem as

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{8.3.78}$$

---

[4] Although this would appear to make $h = 0$, we have merely chosen a temperature scale so that the air temperature is zero and the absolute temperature used in Stefan's law is nonzero.

Table 8.3.1: The First Ten Roots of Equation 8.3.87 and C_n for $hL = 1$

n	α_n	Approximate α_n	C_n
1	2.0288	2.2074	118.9221
2	4.9132	4.9246	31.3414
3	7.9787	7.9813	27.7549
4	11.0855	11.0865	16.2891
5	14.2074	14.2079	14.9916
6	17.3364	17.3366	10.8362
7	20.4692	20.4693	10.2232
8	23.6043	23.6044	8.0999
9	26.7409	26.7410	7.7479
10	29.8786	29.8786	6.4626

with

$$u(x,0) = 100, \quad 0 < x < L, \tag{8.3.79}$$

$$u(0,t) = 0, \quad 0 < t, \tag{8.3.80}$$

and

$$\frac{\partial u(L,t)}{\partial x} + hu(L,t) = 0, \quad 0 < t. \tag{8.3.81}$$

Once again, we assume a product solution $u(x,t) = X(x)T(t)$ with a negative separation constant so that

$$\frac{X''}{X} = \frac{T'}{a^2 T} = -k^2. \tag{8.3.82}$$

We obtain for the x-dependence that

$$X'' + k^2 X = 0, \tag{8.3.83}$$

but the boundary conditions are now

$$X(0) = 0, \quad \text{and} \quad X'(L) + hX(L) = 0. \tag{8.3.84}$$

The most general solution of Equation 8.3.83 is

$$X(x) = A\cos(kx) + B\sin(kx). \tag{8.3.85}$$

However, $A = 0$, because $X(0) = 0$. On the other hand,

$$k\cos(kL) + h\sin(kL) = kL\cos(kL) + hL\sin(kL) = 0, \tag{8.3.86}$$

if $B \neq 0$. The nondimensional number hL is the *Biot number* and depends completely upon the physical characteristics of the rod.

In Chapter 6 we saw how to find the roots of the transcendental equation

$$\alpha + hL\tan(\alpha) = 0, \tag{8.3.87}$$

where $\alpha = kL$. Consequently, if α_n is the nth root of Equation 8.3.87, then the eigenfunction is

$$X_n(x) = \sin(\alpha_n x/L). \tag{8.3.88}$$

In Table 8.3.1, we list the first ten roots of Equation 8.3.87 for $hL = 1$.

In general, we must solve Equation 8.3.87 either numerically or graphically. If α is large, however, we can find approximate values[5] by noting that

$$\cot(\alpha) = -hL/\alpha \approx 0, \tag{8.3.89}$$

or

$$\alpha_n = (2n-1)\pi/2, \tag{8.3.90}$$

where $n = 1, 2, 3, \ldots$. We can obtain a better approximation by setting

$$\alpha_n = (2n-1)\pi/2 - \epsilon_n, \tag{8.3.91}$$

where $\epsilon_n \ll 1$. Substituting into Equation 8.3.89,

$$[(2n-1)\pi/2 - \epsilon_n]\cot[(2n-1)\pi/2 - \epsilon_n] + hL = 0. \tag{8.3.92}$$

We can simplify Equation 8.3.92 to

$$\epsilon_n^2 + (2n-1)\pi\epsilon_n/2 + hL = 0, \tag{8.3.93}$$

because $\cot[(2n-1)\pi/2 - \theta] = \tan(\theta)$, and $\tan(\theta) \approx \theta$ for $\theta \ll 1$. Solving for ϵ_n,

$$\epsilon_n \approx -\frac{2hL}{(2n-1)\pi}, \tag{8.3.94}$$

and

$$\alpha_n \approx \frac{(2n-1)\pi}{2} + \frac{2hL}{(2n-1)\pi}. \tag{8.3.95}$$

In Table 8.3.1 we compare the approximate roots given by Equation 8.3.95 with the actual roots.

The temporal part equals

$$T_n(t) = C_n \exp\left(-k_n^2 a^2 t\right) = C_n \exp\left(-\frac{\alpha_n^2 a^2 t}{L^2}\right). \tag{8.3.96}$$

Consequently, the general solution is

$$u(x,t) = \sum_{n=1}^{\infty} C_n \sin\left(\frac{\alpha_n x}{L}\right) \exp\left(-\frac{\alpha_n^2 a^2 t}{L^2}\right), \tag{8.3.97}$$

---

[5] Using the same technique, Stevens and Luck [Stevens, J. W., and R. Luck, 1999: Explicit approximations for all eigenvalues of the 1-D transient heat conduction equations. *Heat Transfer Eng.*, **20(2)**, 35–41] have found approximate solutions to $\zeta_n \tan(\zeta_n) = Bi$. They showed that

$$\zeta_n \approx z_n + \frac{-B + \sqrt{B^2 - 4C}}{2},$$

where

$$B = z_n + (1 + Bi)\tan(z_n), \qquad C = Bi - z_n \tan(z_n),$$

$$z_n = c_n + \frac{\pi}{4}\left(\frac{Bi - c_n}{Bi + c_n}\right), \qquad c_n = \left(n - \frac{3}{4}\right)\pi.$$

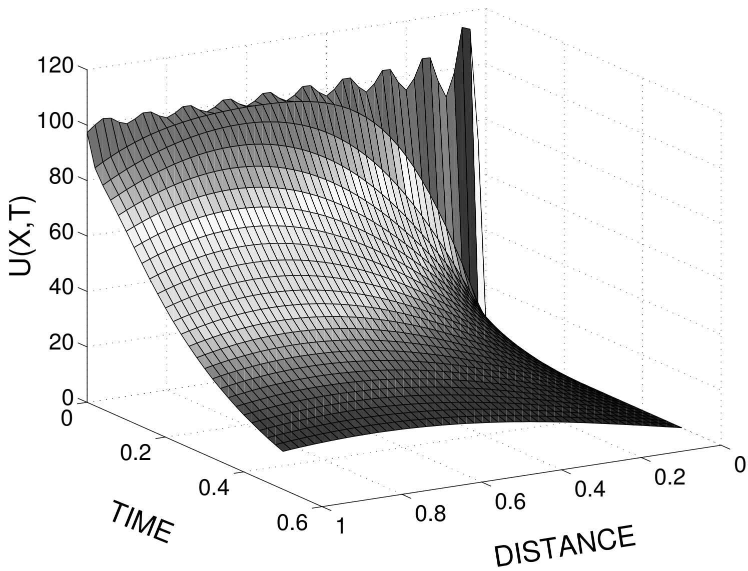


Figure 8.3.5: The temperature $u(x,t)$ within a thin bar as a function of position x/L and time a^2t/L^2 when we allow the bar to radiatively cool at $x = L$ while the temperature is zero at $x = 0$. Initially the temperature was 100.

where α_n is the nth root of Equation 8.3.87.

To determine C_n, we use the initial condition, Equation 8.3.79, and find that

$$100 = \sum_{n=1}^{\infty} C_n \sin\left(\frac{\alpha_n x}{L}\right). \tag{8.3.98}$$

Equation 8.3.98 is an eigenfunction expansion of 100 employing the eigenfunctions from the Sturm-Liouville problem

$$X'' + k^2 X = 0, \tag{8.3.99}$$

and

$$X(0) = X'(L) + hX(L) = 0. \tag{8.3.100}$$

Thus, the coefficient C_n is given by Equation 6.3.4 or

$$C_n = \frac{\int_0^L 100 \sin(\alpha_n x/L)\,dx}{\int_0^L \sin^2(\alpha_n x/L)\,dx}, \tag{8.3.101}$$

as $r(x) = 1$. Performing the integrations,

$$C_n = \frac{100L[1-\cos(\alpha_n)]/\alpha_n}{\frac{1}{2}[L - L\sin(2\alpha_n)/(2\alpha_n)]} = \frac{200[1-\cos(\alpha_n)]}{\alpha_n[1+\cos^2(\alpha_n)/(hL)]}, \tag{8.3.102}$$

because $\sin(2\alpha_n) = 2\cos(\alpha_n)\sin(\alpha_n)$, and $\alpha_n = -hL\tan(\alpha_n)$. The complete solution is

$$u(x,t) = \sum_{n=1}^{\infty} \frac{200[1-\cos(\alpha_n)]}{\alpha_n[1+\cos^2(\alpha_n)/(hL)]} \sin\left(\frac{\alpha_n x}{L}\right) \exp\left(-\frac{\alpha_n^2 a^2 t}{L^2}\right). \tag{8.3.103}$$

Figure 8.3.5 illustrates this solution for $hL = 1$ at various times and positions. It was generated using the MATLAB script

```
clear
hL = 1; M = 200; dx = 0.02; dt = 0.02;
% create initial guess at alpha_n
zero = zeros(M,1);
for n = 1:M
  temp = (2*n-1)*pi; zero(n) = 0.5*temp + 2*hL/temp;
end;
% use Newton-Raphson method to improve values of alpha_n
for n = 1:M; for k = 1:10
  f = zero(n) + hL * tan(zero(n)); fp =1 + hL * sec(zero(n))^2;
  zero(n) = zero(n) - f / fp;
end; end;
% compute Fourier coefficients
for m = 1:M
  a(m) = 200*(1-cos(zero(m)))/(zero(m)*(1+cos(zero(m))^2/hL));
end
% compute grid and initialize solution
X = [0:dx:1]; T = [0:dt:0.5];
u = zeros(length(T),length(X));
XX = repmat(X,[length(T) 1]);
TT = repmat(T',[1 length(X)]);
% compute solution from Equation 8.3.103
for m = 1:M
  u = u + a(m) * sin(zero(m)*XX) .* exp(-zero(m)*zero(m)*TT);
end
surf(XX,TT,u)
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('U(X,T)','Fontsize',20)
```

It is similar to Example 8.3.1 in that the heat lost to the environment occurs either because the temperature at an end is zero or because it radiates heat to space that has the temperature of zero. □

• Example 8.3.6: Refrigeration of apples

Some decades ago, shiploads of apples, going from Australia to England, deteriorated from a disease called "brown heart," which occurred under insufficient cooling conditions. Apples, when placed on shipboard, are usually warm and must be cooled to be carried in cold storage. They also generate heat by their respiration. It was suspected that this heat generation effectively counteracted the refrigeration of the apples, resulting in the "brown heart."

This was the problem that induced Awberry[6] to study the heat distribution within a sphere in which heat is being generated. Awberry first assumed that the apples are initially at a uniform temperature. We can take this temperature to be zero by the appropriate choice of temperature scale. At time $t = 0$, the skins of the apples assume the temperature θ immediately when we introduce them into the hold.

---

[6] Awberry, J. H., 1927: The flow of heat in a body generating heat. *Philos. Mag., Ser. 7*, **4**, 629–638.

Because of the spherical geometry, the nonhomogeneous heat equation becomes

$$\frac{1}{a^2}\frac{\partial u}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{G}{\kappa}, \qquad 0 \le r < b, \quad 0 < t, \tag{8.3.104}$$

where a^2 is the thermal diffusivity, b is the radius of the apple, κ is the thermal conductivity, and G is the heating rate (per unit time per unit volume).

If we try to use separation of variables on Equation 8.3.104, we find that it does not work because of the G/κ term. To circumvent this difficulty, we ask the simpler question of what happens after a very long time. We anticipate that a balance will eventually be established where conduction transports the heat produced within the apple to the surface of the apple where the surroundings absorb it. Consequently, just as we introduced a steady-state solution in Example 8.3.3, we again anticipate a steady-state solution $w(r)$ where the heat conduction removes the heat generated within the apples. The ordinary differential equation

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dw}{dr}\right) = -\frac{G}{\kappa} \tag{8.3.105}$$

gives the steady-state. Furthermore, just as we introduced a transient solution that allowed our solution to satisfy the initial condition, we must also have one here, and the governing equation is

$$\frac{\partial v}{\partial t} = \frac{a^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial v}{\partial r}\right). \tag{8.3.106}$$

Solving Equation 8.3.106 first,

$$w(r) = C + \frac{D}{r} - \frac{Gr^2}{6\kappa}. \tag{8.3.107}$$

The constant D equals zero because the solution must be finite at $r = 0$. Since the steady-state solution must satisfy the boundary condition $w(b) = \theta$,

$$C = \theta + \frac{Gb^2}{6\kappa}. \tag{8.3.108}$$

Turning to the transient problem, we introduce a new dependent variable $y(r,t) = rv(r,t)$. This new dependent variable allows us to replace Equation 8.3.106 with

$$\frac{\partial y}{\partial t} = a^2\frac{\partial^2 y}{\partial r^2}, \tag{8.3.109}$$

which we can solve. If we assume that $y(r,t) = R(r)T(t)$ and we only have a negative separation constant, the $R(r)$ equation becomes

$$\frac{d^2R}{dr^2} + k^2 R = 0, \tag{8.3.110}$$

which has the solution

$$R(r) = A\cos(kr) + B\sin(kr). \tag{8.3.111}$$

The constant A equals zero because the solution, Equation 8.3.111, must vanish at $r = 0$ so that $v(0,t)$ remains finite. However, because $\theta = w(b) + v(b,t)$ for all time and $v(b,t) = R(b)T(t)/b = 0$, then $R(b) = 0$. Consequently, $k_n = n\pi/b$, and

$$v_n(r,t) = \frac{B_n}{r}\sin\left(\frac{n\pi r}{b}\right)\exp\left(-\frac{n^2\pi^2a^2t}{b^2}\right). \tag{8.3.112}$$

Superposition gives the total solution, which equals

$$u(r,t) = \theta + \frac{G}{6\kappa}(b^2 - r^2) + \sum_{n=1}^{\infty} \frac{B_n}{r} \sin\left(\frac{n\pi r}{b}\right) \exp\left(-\frac{n^2\pi^2a^2t}{b^2}\right). \tag{8.3.113}$$

Finally, we determine the coefficients B_n by the initial condition that $u(r,0) = 0$. Therefore,

$$B_n = -\frac{2}{b}\int_0^b r\left[\theta + \frac{G}{6\kappa}(b^2 - r^2)\right] \sin\left(\frac{n\pi r}{b}\right)\, dr = \frac{2\theta b}{n\pi}(-1)^n + \frac{2G}{\kappa}\left(\frac{b}{n\pi}\right)^3(-1)^n. \tag{8.3.114}$$

The complete solution is

$$u(r,t) = \theta + \frac{2\theta b}{r\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin\left(\frac{n\pi r}{b}\right)\exp\left(-\frac{n^2\pi^2a^2t}{b^2}\right) \tag{8.3.115}$$
$$+ \frac{G}{6\kappa}(b^2 - r^2) + \frac{2Gb^3}{r\kappa\pi^3}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^3}\sin\left(\frac{n\pi r}{b}\right)\exp\left(-\frac{n^2\pi^2a^2t}{b^2}\right).$$

The first line of Equation 8.3.115 gives the temperature distribution due to the imposition of the temperature θ on the surface of the apple while the second line gives the rise in the temperature due to the interior heating.

Returning to our original problem of whether the interior heating is strong enough to counteract the cooling by refrigeration, we merely use the second line of Equation 8.3.115 to find how much the temperature deviates from what we normally expect. Because the highest temperature exists at the center of each apple, its value there is the only one of interest in this problem. Assuming $b = 4$ cm as the radius of the apple, $a^2G/\kappa = 1.33\times10^{-5}$ °C/s, and $a^2 = 1.55 \times 10^{-3}$ cm^2/s, the temperature effect of the heat generation is very small, only 0.0232 °C when, after about 2 hours, the temperatures within the apples reach equilibrium. Thus, we must conclude that heat generation within the apples is not the cause of brown heart.

We now know that brown heart results from an excessive concentration of carbon dioxide and an insufficient amount of oxygen in the storage hold.[7] Presumably this atmosphere affects the metabolic activities that are occurring in the apple[8] and leads to low-temperature breakdown. □

• Example 8.3.7

In this example we illustrate how separation of variables can be employed in solving the axisymmetric heat equation in an infinitely long cylinder. In circular coordinates the heat equation is

$$\frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right), \qquad 0 \le r < b, \quad 0 < t, \tag{8.3.116}$$

---

[7] Thornton, N. C., 1931: The effect of carbon dioxide on fruits and vegetables in storage. *Contrib. Boyce Thompson Inst.*, **3**, 219–244.

[8] Fidler, J. C., and C. J. North, 1968: The effect of conditions of storage on the respiration of apples. IV. Changes in concentration of possible substrates of respiration, as related to production of carbon dioxide and uptake of oxygen by apples at low temperatures. *J. Hortic. Sci.*, **43**, 429–439.

where r denotes the radial distance and a^2 denotes the thermal diffusivity. Let us assume that we heated this cylinder of radius b to the uniform temperature T_0 and then allowed it to cool by having its surface held at the temperature of zero starting from the time $t = 0$.

We begin by assuming that the solution is of the form $u(r,t) = R(r)T(t)$ so that

$$\frac{1}{R}\left(\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr}\right) = \frac{1}{a^2T}\frac{dT}{dt} = -\frac{k^2}{b^2}. \qquad \textbf{(8.3.117)}$$

The only values of the separation constant that yield nontrivial solutions are negative. The nontrivial solutions are $R(r) = J_0(kr/b)$, where J_0 is the Bessel function of the first kind and zeroth order. A separation constant of zero gives $R(r) = \ln(r)$, which becomes infinite at the origin. Positive separation constants yield the modified Bessel function $I_0(kr/b)$. Although this function is finite at the origin, it cannot satisfy the boundary condition that $u(b,t) = R(b)T(t) = 0$, or $R(b) = 0$.

The boundary condition that $R(b) = 0$ requires that $J_0(k) = 0$. This transcendental equation yields an infinite number of constants k_n. For each k_n, the temporal part of the solution satisfies the differential equation

$$\frac{dT_n}{dt} + \frac{k_n^2a^2}{b^2}T_n = 0, \qquad \textbf{(8.3.118)}$$

which has the solution

$$T_n(t) = A_n \exp\left(-\frac{k_n^2a^2}{b^2}t\right). \qquad \textbf{(8.3.119)}$$

Consequently, the product solutions are

$$u_n(r,t) = A_nJ_0\left(k_n\frac{r}{b}\right)\exp\left(-\frac{k_n^2a^2}{b^2}t\right). \qquad \textbf{(8.3.120)}$$

The total solution is a linear superposition of all of the particular solutions or

$$u(r,t) = \sum_{n=1}^{\infty} A_nJ_0\left(k_n\frac{r}{b}\right)\exp\left(-\frac{k_n^2a^2}{b^2}t\right). \qquad \textbf{(8.3.121)}$$

Our final task remains to determine A_n. From the initial condition that $u(r,0) = T_0$,

$$u(r,0) = T_0 = \sum_{n=1}^{\infty} A_nJ_0\left(k_n\frac{r}{b}\right). \qquad \textbf{(8.3.122)}$$

From Equation 6.5.38 and Equation 6.5.45,

$$A_n = \frac{2T_0}{J_1^2(k_n)b^2}\int_0^b rJ_0\left(k_n\frac{r}{b}\right)dr = \frac{2T_0}{k_n^2J_1^2(k_n)}\left(\frac{k_nr}{b}\right)J_1\left(k_n\frac{r}{b}\right)\bigg|_0^b = \frac{2T_0}{k_nJ_1(k_n)} \qquad \textbf{(8.3.123)}$$

from Equation 6.5.28. Thus, the complete solution is

$$u(r,t) = 2T_0\sum_{n=1}^{\infty}\frac{1}{k_nJ_1(k_n)}J_0\left(k_n\frac{r}{b}\right)\exp\left(-\frac{k_n^2a^2}{b^2}t\right). \qquad \textbf{(8.3.124)}$$

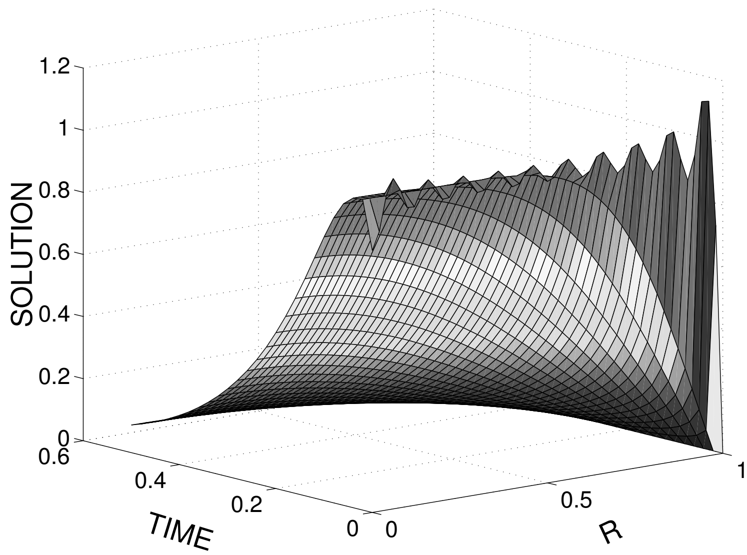


Figure 8.3.6: The temperature $u(r,t)/T_0$ within an infinitely long cylinder at various positions r/b and times a^2t/b^2 that we initially heated to the uniform temperature T_0 and then allowed to cool by forcing its surface to equal zero.

Figure 8.3.6 illustrates the solution, Equation 8.3.124, for various Fourier numbers a^2t/b^2. It was generated using the MATLAB script

```
clear
M = 20; dr = 0.02; dt = 0.02;
% load in zeros of J_0
zero( 1) =  2.40482; zero( 2) =  5.52007; zero( 3) =  8.65372;
zero( 4) = 11.79153; zero( 5) = 14.93091; zero( 6) = 18.07106;
zero( 7) = 21.21164; zero( 8) = 24.35247; zero( 9) = 27.49347;
zero(10) = 30.63461; zero(11) = 33.77582; zero(12) = 36.91710;
zero(13) = 40.05843; zero(14) = 43.19979; zero(15) = 46.34119;
zero(16) = 49.48261; zero(17) = 52.62405; zero(18) = 55.76551;
zero(19) = 58.90698; zero(20) = 62.04847;
% compute Fourier coefficients
for m = 1:M
  a(m) = 2 / (zero(m)*besselj(1,zero(m)));
end
% compute grid and initialize solution
R = [0:dr:1]; T = [0:dt:0.5];
u = zeros(length(T),length(R));
RR = repmat(R,[length(T) 1]);
TT = repmat(T',[1 length(R)]);
% compute solution from Equation 8.3.124
for m = 1:M
  u = u + a(m)*besselj(0,zero(m)*RR).*exp(-zero(m)*zero(m)*TT);
end
surf(RR,TT,u)
xlabel('R','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

It is similar to Example 8.3.1 except that we are in cylindrical coordinates. Heat flows from the interior and is removed at the cylinder's surface where the temperature equals zero. The initial oscillations of the solution result from Gibbs phenomena because we have a jump in the temperature field at $r = b$. □

• Example 8.3.8

In this example[9] we find the evolution of the temperature field within a cylinder of radius b as it radiatively cools from an initial uniform temperature T_0. The heat equation is

$$\frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right), \qquad 0 \le r < b, \quad 0 < t, \tag{8.3.125}$$

which we will solve by separation of variables $u(r,t) = R(r)T(t)$. Therefore,

$$\frac{1}{R}\left(\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr}\right) = \frac{1}{a^2T}\frac{dT}{dt} = -\frac{k^2}{b^2}, \tag{8.3.126}$$

because only a negative separation constant yields an $R(r)$ that is finite at the origin and satisfies the boundary condition. This solution is $R(r) = J_0(kr/b)$, where J_0 is the Bessel function of the first kind and zeroth order.

The radiative boundary condition can be expressed as

$$\frac{\partial u(b,t)}{\partial r} + hu(b,t) = T(t)\left[\frac{dR(b)}{dr} + hR(b)\right] = 0. \tag{8.3.127}$$

Because $T(t) \neq 0$,

$$kJ_0'(k) + hbJ_0(k) = -kJ_1(k) + hbJ_0(k) = 0, \tag{8.3.128}$$

where the product hb is the Biot number. The solution of the transcendental equation, Equation 8.3.128, yields an infinite number of distinct constants k_n. For each k_n, the temporal part equals the solution of

$$\frac{dT_n}{dt} + \frac{k_n^2a^2}{b^2}T_n = 0, \tag{8.3.129}$$

or

$$T_n(t) = A_n \exp\left(-\frac{k_n^2a^2}{b^2}t\right). \tag{8.3.130}$$

The product solution is, therefore,

$$u_n(r,t) = A_nJ_0\left(k_n\frac{r}{b}\right)\exp\left(-\frac{k_n^2a^2}{b^2}t\right) \tag{8.3.131}$$

and the most general solution is a sum of these product solutions

$$u(r,t) = \sum_{n=1}^{\infty} A_nJ_0\left(k_n\frac{r}{b}\right)\exp\left(-\frac{k_n^2a^2}{b^2}t\right). \tag{8.3.132}$$

---

[9] For another example of solving the heat equation with Robin boundary conditions, see Section 3.2 in Balakotaiah, V., N. Gupta, and D. H. West, 2000: A simplified model for analyzing catalytic reactions in short monoliths. *Chem. Eng. Sci.*, **55**, 5367–5383.

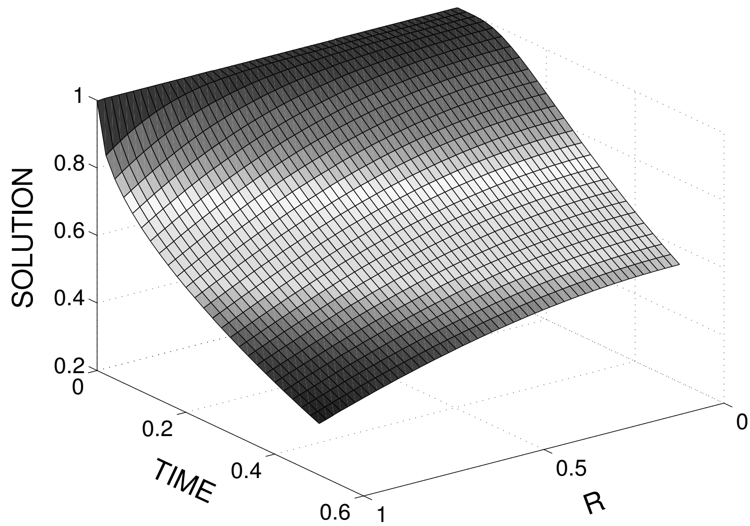


Figure 8.3.7: The temperature $u(r,t)/T_0$ within an infinitely long cylinder at various positions r/b and times a^2t/b^2 that we initially heated to the temperature T_0 and then allowed to radiatively cool with $hb = 1$.

Finally, we must determine A_n. From the initial condition that $u(r,0) = T_0$,

$$u(r,0) = T_0 = \sum_{n=1}^{\infty} A_n J_0\left(k_n \frac{r}{b}\right), \tag{8.3.133}$$

where

$$A_n = \frac{2k_n^2 T_0}{b^2[k_n^2 + b^2h^2]J_0^2(k_n)} \int_0^b r J_0\left(k_n \frac{r}{b}\right)\, dr = \frac{2T_0}{[k_n^2 + b^2h^2]J_0^2(k_n)} \left(\frac{k_n r}{b}\right) J_1\left(k_n \frac{r}{b}\right)\Big|_0^b \tag{8.3.134}$$

$$= \frac{2k_n T_0 J_1(k_n)}{[k_n^2 + b^2h^2]J_0^2(k_n)} = \frac{2k_n T_0 J_1(k_n)}{k_n^2 J_0^2(k_n) + b^2h^2 J_0^2(k_n)} = \frac{2k_n T_0 J_1(k_n)}{k_n^2 J_0^2(k_n) + k_n^2 J_1^2(k_n)} \tag{8.3.135}$$

$$= \frac{2T_0 J_1(k_n)}{k_n[J_0^2(k_n) + J_1^2(k_n)]}, \tag{8.3.136}$$

which follows from Equation 6.5.28, Equation 6.5.38, Equation 6.5.45, and Equation 8.3.128. Consequently, the complete solution is

$$u(r,t) = 2T_0 \sum_{n=1}^{\infty} \frac{J_1(k_n)}{k_n[J_0^2(k_n) + J_1^2(k_n)]} J_0\left(k_n \frac{r}{b}\right) \exp\left(-\frac{k_n^2 a^2}{b^2} t\right). \tag{8.3.137}$$

Figure 8.3.7 illustrates the solution, Equation 8.3.137, for various Fourier numbers a^2t/b^2 with $hb = 1$. It was created using the MATLAB script

```
clear
hb = 1; m=0; M = 100; dr = 0.02; dt = 0.02;
% find k_n which satisfies hb J_0(k) = k J_1(k)
for n = 1:10000
```

```
  k1 = 0.05*n; k2 = 0.05*(n+1);
  y1 = hb * besselj(0,k1) - k1 * besselj(1,k1);
  y2 = hb * besselj(0,k2) - k2 * besselj(1,k2);
  if y1*y2 <= 0; m = m+1; zero(m) = k1; end;
end;
% use Newton-Raphson method to improve values of k_n
for n = 1:M; for k = 1:5
  term0 = besselj(0,zero(n));
  term1 = besselj(1,zero(n));
  term2 = besselj(2,zero(n));
  f = hb * term0 - zero(n) * term1;
  fp = 0.5*zero(n)*(term2-term0) - (1+hb)*term1;
  zero(n) = zero(n) - f / fp;
end; end;
% compute Fourier coefficients
for m = 1:M
  denom = zero(m)*(besselj(0,zero(m))^2+besselj(1,zero(m))^2);
  a(m) = 2 * besselj(1,zero(m)) / denom;
end
% compute grid and initialize solution
R = [0:dr:1]; T = [0:dt:0.5];
u = zeros(length(T),length(R));
RR = repmat(R,[length(T) 1]);
TT = repmat(T',[1 length(R)]);
% compute solution from Equation 8.3.137
for m = 1:M
  u = u + a(m)*besselj(0,zero(m)*RR).*exp(-zero(m)*zero(m)*TT);
end
surf(RR,TT,u)
xlabel('R','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

These results are similar to Example 8.3.5 except that we are in cylindrical coordinates. Heat flows from the interior and is removed at the cylinder's surface where it radiates to space at the temperature zero. Note that we do *not* suffer from Gibbs phenomena in this case because there is no initial jump in the temperature distribution. □

• Example 8.3.9: Temperature within an electrical cable

In the design of cable installations we need the temperature reached within an electrical cable as a function of current and other parameters. To this end,[10] let us solve the nonhomogeneous heat equation in cylindrical coordinates with a radiation boundary condition.

The derivation of the heat equation follows from the conservation of energy:

$$\text{heat generated} = \text{heat dissipated} + \text{heat stored},$$

[10] Iskenderian, H. P., and W. J. Horvath, 1946: Determination of the temperature rise and the maximum safe current through multiconductor electric cables. *J. Appl. Phys.*, **17**, 255–262.

or

$$I^2RN\,dt = -\kappa\left[2\pi r\left.\frac{\partial u}{\partial r}\right|_r - 2\pi(r+\Delta r)\left.\frac{\partial u}{\partial r}\right|_{r+\Delta r}\right]dt + 2\pi r\Delta r c\rho\,du, \qquad (\mathbf{8.3.138})$$

where I is the current through each wire, R is the resistance of each conductor, N is the number of conductors in the shell between radii r and $r+\Delta r = 2\pi m r\Delta r/(\pi b^2)$, b is the radius of the cable, m is the total number of conductors in the cable, κ is the thermal conductivity, ρ is the density, c is the average specific heat, and u is the temperature. In the limit of $\Delta r \to 0$, Equation 8.3.138 becomes

$$\frac{\partial u}{\partial t} = A + \frac{a^2}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right), \qquad 0 \le r < b, \quad 0 < t, \qquad (\mathbf{8.3.139})$$

where $A = I^2Rm/(\pi b^2 c\rho)$, and $a^2 = \kappa/(\rho c)$.

Equation 8.3.139 is the nonhomogeneous heat equation for an infinitely long, axisymmetric cylinder. From Example 8.3.3, we know that we must write the temperature as the sum of a steady-state and transient solution: $u(r,t) = w(r) + v(r,t)$. The steady-state solution $w(r)$ satisfies

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dw}{dr}\right) = -\frac{A}{a^2}, \qquad (\mathbf{8.3.140})$$

or

$$w(r) = T_c - \frac{Ar^2}{4a^2}, \qquad (\mathbf{8.3.141})$$

where T_c is the (yet unknown) temperature in the center of the cable.

The transient solution $v(r,t)$ is governed by

$$\frac{\partial v}{\partial t} = a^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v}{\partial r}\right), \qquad 0 \le r < b, \quad 0 < t, \qquad (\mathbf{8.3.142})$$

with the initial condition that $u(r,0) = T_c - Ar^2/(4a^2) + v(0,t) = 0$. At the surface $r = b$, heat radiates to free space so that the boundary condition is $u_r = -hu$, where h is the surface conductance. Because the temperature equals the steady-state solution when all transient effects die away, $w(r)$ must satisfy this radiation boundary condition regardless of the transient solution. This requires that

$$T_c = \frac{A}{a^2}\left(\frac{b^2}{4} + \frac{b}{2h}\right). \qquad (\mathbf{8.3.143})$$

Therefore, $v(r,t)$ must satisfy $v_r(b,t) = -hv(b,t)$ at $r = b$.

We find the transient solution $v(r,t)$ by separation of variables $v(r,t) = R(r)T(t)$. Substituting into Equation 8.3.142,

$$\frac{1}{rR}\frac{d}{dr}\left(r\frac{dR}{dr}\right) = \frac{1}{a^2T}\frac{dT}{dt} = -k^2, \qquad (\mathbf{8.3.144})$$

or

$$\frac{d}{dr}\left(r\frac{dR}{dr}\right) + k^2rR = 0, \qquad (\mathbf{8.3.145})$$

and

$$\frac{dT}{dt} + k^2a^2T = 0, \qquad (\mathbf{8.3.146})$$

with $R'(b) = -hR(b)$. The only solution of Equation 8.3.145 that remains finite at $r = 0$ and satisfies the boundary condition is $R(r) = J_0(kr)$, where J_0 is the zero-order Bessel function of the first kind. Substituting $J_0(kr)$ into the boundary condition, the transcendental equation is

$$kbJ_1(kb) - hbJ_0(kb) = 0. \tag{8.3.147}$$

For a given value of h and b, Equation 8.3.147 yields an infinite number of unique zeros k_n.

The corresponding temporal solution to the problem is

$$T_n(t) = A_n \exp(-a^2 k_n^2 t), \tag{8.3.148}$$

so that the sum of the product solutions is

$$v(r,t) = \sum_{n=1}^{\infty} A_n J_0(k_n r) \exp(-a^2 k_n^2 t). \tag{8.3.149}$$

Our final task remains to compute A_n. By evaluating Equation 8.3.149 at $t = 0$,

$$v(r,0) = \frac{Ar^2}{4a^2} - T_c = \sum_{n=1}^{\infty} A_n J_0(k_n r), \tag{8.3.150}$$

which is a Fourier-Bessel series in $J_0(k_n r)$. In Section 6.5 we showed how to compute the coefficient of a Fourier-Bessel series that is expanded in $J_0(k_n r)$ and that has a boundary condition of the form given Equation 8.3.147. Applying those results here, we have that

$$A_n = \frac{2k_n^2}{(k_n^2 b^2 + h^2 b^2) J_0^2(k_n b)} \int_0^b r \left(\frac{Ar^2}{4a^2} - T_c \right) J_0(k_n r)\, dr \tag{8.3.151}$$

from Equation 6.5.38 and Equation 6.5.47. Carrying out the indicated integrations,

$$A_n = \frac{2}{(k_n^2 + h^2) J_0^2(k_n b)} \left[\left(\frac{Ak_n b}{4a^2} - \frac{A}{k_n b a^2} - \frac{T_c k_n}{b} \right) J_1(k_n b) + \frac{A}{2a^2} J_0(k_n b) \right]. \tag{8.3.152}$$

We obtained Equation 8.3.152 by using Equation 6.5.28 and integrating by parts as shown in Example 6.5.5.

To illustrate this solution, let us compute it for the typical parameters $b = 4$ cm, $hb = 1$, $a^2 = 1.14$ cm^2/s, $A = 2.2747$ °C/s, and $T_c = 23.94$°C. The value of A corresponds to 37 wires of #6 AWG copper wire within a cable carrying a current of 22 amp.

Figure 8.3.8 illustrates the solution as a function of radius at various times. It was created using the MATLAB script

```
clear
asq = 1.14; A = 2.2747; b = 4; dr = 0.02; dt = 0.02;
hb = 1; m=0; M = 10; T_c = 23.94;
const1 = A * b * b / (4 * asq); const2 = A * b * b / asq;
const3 = A * b * b / (2 * asq);
% find k_nb which satisfies hb J_0(kb) = kb J_1(kb)
for n = 1:10000
  k1 = 0.05*n; k2 = 0.05*(n+1);
  y1 = hb * besselj(0,k1) - k1 * besselj(1,k1);
  y2 = hb * besselj(0,k2) - k2 * besselj(1,k2);
```

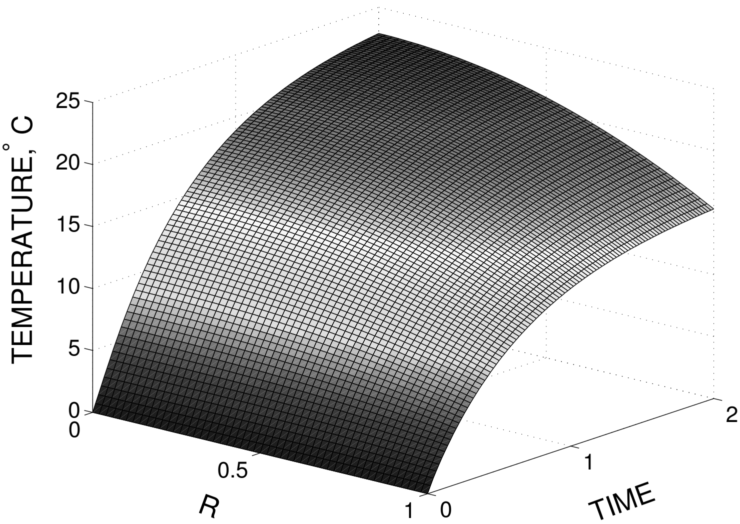


Figure 8.3.8: The temperature field (in degrees Celsius) within an electric copper cable containing 37 wires and a current of 22 amperes at various positions r/b and times a^2t/b^2. Initially the temperature was zero and then we allow the cable to cool radiatively as it is heated. The parameters are $hb = 1$ and the radius of the cable $b = 4$ cm.

```
  if y1*y2 <= 0; m = m+1; zero(m) = k1; end;
end;
% use Newton-Raphson method to improve values of k_n
for n = 1:M; for k = 1:5
  term0 = besselj(0,zero(n));
  term1 = besselj(1,zero(n));
  term2 = besselj(2,zero(n));
  f = hb * term0 - zero(n) * term1;
  fp = 0.5*zero(n)*(term2-term0) - (1+hb)*term1;
  zero(n) = zero(n) - f / fp;
end; end;
for m = 1:M
  denom = (zero(m)*zero(m)+hb*hb)*besselj(0,zero(m))^2;
  a(m) = ((const1-T_c)*zero(m) ...
       - const2/zero(m))*besselj(1,zero(m)) ...
       + const3 * besselj(0,zero(m));
  a(m) = 2 * a(m) / denom;
end
% compute grid and initialize solution
R = [0:dr:1]; T = [0:dt:2];
u = T_c * ones(length(T),length(R));
RR = repmat(R,[length(T) 1]);
TT = repmat(T',[1 length(R)]);
% compute solution u(r,t) = w(r) + v(r,t)
u = u - const1 * RR .* RR;
for m = 1:M
  u = u + a(m)*besselj(0,zero(m)*RR).*exp(-zero(m)*zero(m)*TT);
end
```

```
surf(RR,TT,u); axis([0 1 0 2 0 25]);
xlabel('R','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('TEMPERATURE,^\circ C','Fontsize',20)
```

From an initial temperature of zero, the temperature rises due to the constant electrical heating. After a short period of time, it reaches its steady-state distribution given by Equation 8.3.140. The cable is coolest at the surface where heat is radiating away. Heat flows from the interior to replace the heat lost by radiation.

Problems

For Problems 1–5, solve the heat equation $u_t = a^2 u_{xx}$, $0 < x < \pi$, $0 < t$, subject to the boundary conditions that $u(0,t) = u(\pi,t) = 0$, $0 < t$, and the following initial conditions for $0 < x < \pi$. Then plot your results using MATLAB.

1. $u(x,0) = A$, a constant

2. $u(x,0) = \sin^3(x) = [3\sin(x) - \sin(3x)]/4$

3. $u(x,0) = x$

4. $u(x,0) = \pi - x$

5. $u(x,0) = \begin{cases} x, & 0 < x < \pi/2, \\ \pi - x, & \pi/2 < x < \pi. \end{cases}$

For Problems 6–10, solve the heat equation $u_t = a^2 u_{xx}$, $0 < x < \pi$, $0 < t$, subject to the boundary conditions that $u_x(0,t) = u_x(\pi,t) = 0$, $0 < t$, and the following initial conditions for $0 < x < \pi$. Then plot your results using MATLAB.

6. $u(x,0) = 1$

7. $u(x,0) = x$

8. $u(x,0) = \cos^2(x) = [1 + \cos(2x)]/2$

9. $u(x,0) = \pi - x$

10. $u(x,0) = \begin{cases} T_0, & 0 < x < \pi/2, \\ T_1, & \pi/2 < x < \pi. \end{cases}$

For Problems 11–17, solve the heat equation $u_t = a^2 u_{xx}$, $0 < x < \pi$, $0 < t$, subject to the following boundary conditions and initial condition. Then plot your results using MATLAB.

11. $u_x(0,t) = u(\pi,t) = 0$, $0 < t$; $u(x,0) = x^2 - \pi^2$, $0 < x < \pi$

12. $u(0,t) = u(\pi,t) = T_0$, $0 < t$; $u(x,0) = T_1 \neq T_0$, $0 < x < \pi$

13. $u(0,t) = 0$, $u_x(\pi,t) = 0$, $0 < t$; $u(x,0) = 1$, $0 < x < \pi$

14. $u(0,t) = 0$, $u_x(\pi,t) = 0$, $0 < t$; $u(x,0) = x$, $0 < x < \pi$

15. $u(0,t) = 0$, $u_x(\pi,t) = 0$, $0 < t$; $u(x,0) = \pi - x$, $0 < x < \pi$

16. $u(0,t) = T_0$, $u_x(\pi,t) = 0$, $0 < t$; $u(x,0) = T_1 \neq T_0$, $0 < x < \pi$

17. $u(0,t) = 0, u(\pi,t) = T_0$, $0 < t$; $u(x,0) = T_0$, $0 < x < \pi$

18. It is well known that a room with masonry walls is often very difficult to heat. Consider a wall of thickness L, conductivity κ, and diffusivity a^2, which we heat by a surface heat flux at a constant rate H. The temperature of the outside (out-of-doors) face of the wall remains constant at T_0 and the entire wall initially has the uniform temperature T_0. Let us find the temperature of the inside face as a function of time.[11]

We begin by solving the heat conduction problem

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

subject to the boundary conditions that

$$\frac{\partial u(0,t)}{\partial x} = -\frac{H}{\kappa}, \qquad \text{and} \qquad u(L,t) = T_0,$$

and the initial condition that $u(x,0) = T_0$. Show that the temperature field equals

$$u(x,t) = T_0 + \frac{HL}{\kappa}\left\{1 - \frac{x}{L} - \frac{8}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}\cos\left[\frac{(2n-1)\pi x}{2L}\right]\exp\left[-\frac{(2n-1)^2\pi^2 a^2 t}{4L^2}\right]\right\}.$$

Therefore, the rise of temperature at the interior wall $x = 0$ is

$$\frac{HL}{\kappa}\left\{1 - \frac{8}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}\exp\left[-\frac{(2n-1)^2\pi^2 a^2 t}{4L^2}\right]\right\},$$

or

$$\frac{8HL}{\kappa\pi^2}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}\left\{1 - \exp\left[-\frac{(2n-1)^2\pi^2 a^2 t}{4L^2}\right]\right\}.$$

For $a^2t/L^2 \leq 1$, this last expression can be approximated[12] by $2Hat^{1/2}/\ \pi^{1/2}\kappa$. We thus see that the temperature will initially rise as the square root of time and diffusivity and

---

[11] See Dufton, A. F., 1927: The warming of walls. *Philos. Mag., Ser. 7*, **4**, 888–889.

[12] Let us define the function:

$$f(t) = \sum_{n=1}^{\infty}\frac{1 - \exp[-(2n-1)^2\pi^2 a^2 t/L^2]}{(2n-1)^2}.$$

Then

$$f'(t) = \frac{a^2\pi^2}{L^2}\sum_{n=1}^{\infty}\exp[-(2n-1)^2\pi^2 a^2 t/L^2].$$

Consider now the integral

$$\int_0^{\infty}\exp\left(-\frac{a^2\pi^2 t}{L^2}x^2\right)\,dx = \frac{L}{2a\sqrt{\pi t}}.$$

If we approximate this integral by using the trapezoidal rule with $\Delta x = 2$, then

$$\int_0^{\infty}\exp\left(-\frac{a^2\pi^2 t}{L^2}x^2\right)\,dx \approx 2\sum_{n=1}^{\infty}\exp[-(2n-1)^2\pi^2 a^2 t/L^2],$$

and $f'(t) \approx a\pi^{3/2}/(4Lt^{1/2})$. Integrating and using $f(0) = 0$, we finally have $f(t) \approx a\pi^{3/2}t^{1/2}/(2L)$. The smaller a^2t/L^2 is, the smaller the error will be. For example, if $t = L^2/a^2$, then the error is 2.4%.

inversely with conductivity. For an average rock, $\kappa = 0.0042$ g/cm-s, and $a^2 = 0.0118$ cm^2/s, while for wood (spruce) $\kappa = 0.0003$ g/cm-s, and $a^2 = 0.0024$ cm^2/s.

The same set of equations applies to heat transfer within a transistor operating at low frequencies.[13] At the junction ($x = 0$) heat is produced at the rate of H and flows to the transistor's supports ($x = \pm L$) where it is removed. The supports are maintained at the temperature T_0, which is also the initial temperature of the transistor.

19. The linearized Boussinesq equation[14]

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

governs the height of the water table $u(x,t)$ above some reference point, where a^2 is the product of the storage coefficient times the hydraulic coefficient divided by the aquifer thickness. A typical value of a^2 is 10 m^2/min. Consider the problem of a strip of land of width L that separates two reservoirs of depth h_1. Initially the height of the water table would be h_1. Suddenly we lower the reservoir on the right $x = L$ to a depth h_2 [$u(0,t) = h_1$, $u(L,t) = h_2$, and $u(x,0) = h_1$]. Find the height of the water table at any position x within the aquifer and any time $t > 0$.

20. The equation (see Problem 19)

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

governs the height of the water table $u(x,t)$. Consider the problem[15] of a piece of land that suddenly has two drains placed at the points $x = 0$ and $x = L$ so that $u(0,t) = u(L,t) = 0$. If the water table initially has the profile $u(x,0) = 8H(L^3x - 3L^2x^2 + 4Lx^3 - 2x^4)/L^4$, find the height of the water table at any point within the aquifer and any time $t > 0$.

21. We want to find the rise of the water table of an aquifer, which we sandwich between a canal and impervious rocks if we suddenly raise the water level in the canal h_0 units above its initial elevation and then maintain the canal at this level. The linearized Boussinesq equation (see Problem 19)

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

governs the level of the water table with the boundary conditions $u(0,t) = h_0$, and $u_x(L,t) = 0$, and the initial condition $u(x,0) = 0$. Find the height of the water table at any point in the aquifer and any time $t > 0$.

22. Solve the nonhomogeneous heat equation

$$\frac{\partial u}{\partial t} - a^2\frac{\partial^2 u}{\partial x^2} = e^{-x}, \qquad 0 < x < \pi, \quad 0 < t,$$

[13] Mortenson, K. E., 1957: Transistor junction temperature as a function of time. *Proc. IRE*, **45**, 504–513. Equation 2a should read $T_x = -F/k$.

[14] See, for example, Van Schilfgaarde, J., 1970: Theory of flow to drains. *Advances in Hydroscience*, No. 6, Academic Press, 81–85.

[15] For a similar problem, see Dumm, L. D., 1954: New formula for determining depth and spacing of subsurface drains in irrigated lands. *Agric. Eng.*, **35**, 726–730.

subject to the boundary conditions $u(0,t) = u_x(\pi,t) = 0$, $0 < t$, and the initial condition $u(x,0) = f(x)$, $0 < x < \pi$.

23. Solve the nonhomogeneous heat equation

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = -1, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions $u_x(0,t) = u_x(1,t) = 0$, $0 < t$, and the initial condition $u(x,0) = \frac{1}{2}(1 - x^2)$, $0 < x < 1$. Hint: Note that any function of time satisfies the boundary conditions.

24. Solve the nonhomogeneous heat equation

$$\frac{\partial u}{\partial t} - a^2 \frac{\partial^2 u}{\partial x^2} = A\cos(\omega t), \qquad 0 < x < \pi, \quad 0 < t,$$

subject to the boundary conditions $u_x(0,t) = u_x(\pi,t) = 0$, $0 < t$, and the initial condition $u(x,0) = f(x)$, $0 < x < \pi$. Hint: Note that any function of time satisfies the boundary conditions.

25. Solve the nonhomogeneous heat equation

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = \begin{cases} x, & 0 < x \le \pi/2, \\ \pi - x, & \pi/2 \le x < \pi, \end{cases} \qquad 0 < x < \pi, \quad 0 < t,$$

subject to the boundary conditions $u(0,t) = u(\pi,t) = 0$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < \pi$. Hint: Represent the forcing function as a half-range Fourier sine expansion over the interval $(0,\pi)$.

26. A uniform conducting rod of length L and thermometric diffusivity a^2 is initially at temperature zero. We supply heat uniformly throughout the rod so that the heat conduction equation is

$$a^2 \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} - P, \qquad 0 < x < L, \quad 0 < t,$$

where P is the rate at which the temperature would rise if there was no conduction. If we maintain the ends of the rod at the temperature of zero, find the temperature at any position and subsequent time. How would the solution change if the boundary conditions became $u(0,t) = u(L,t) = A \neq 0$, $0 < t$, and the initial conditions read $u(x,0) = A$, $0 < x < L$?

27. Solve the nonhomogeneous heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + \frac{A_0}{c\rho}, \qquad 0 < x < L, \quad 0 < t,$$

where $a^2 = \kappa/c\rho$, with the boundary conditions that

$$\frac{\partial u(0,t)}{\partial x} = 0, \quad \kappa \frac{\partial u(L,t)}{\partial x} + hu(L,t) = 0, \quad 0 < t,$$

and the initial condition that $u(x,0) = 0$, $0 < x < L$.

28. Find the solution of

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - u, \qquad 0 < x < L, \quad 0 < t,$$

with the boundary conditions $u(0,t) = 1$, and $u(L,t) = 0$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < L$.

29. Solve[16]

$$\frac{\partial u}{\partial t} + k_1 u = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

with the boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = 0, \quad a^2 \frac{\partial u(L,t)}{\partial x} + k_2 u(L,t) = 0, \quad 0 < t,$$

and the initial condition $u(x,0) = u_0$, $0 < x < L$.

30. Solve

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

with the boundary conditions

$$a^2 \frac{\partial u(0,t)}{\partial x} = u(0,t), \quad \frac{\partial u(1,t)}{\partial x} = 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = 1$, $0 < x < 1$. Hint: Let $u(x,t) = v(x,t) \exp[(2x - t)/(4a^2)]$ so that the problem becomes

$$\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 v}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

with the boundary conditions

$$2a^2 \frac{\partial v(0,t)}{\partial x} = v(0,t), \quad 2a^2 \frac{\partial v(1,t)}{\partial x} = -v(1,t), \qquad 0 < t,$$

and the initial condition $v(x,0) = \exp[-x/(2a^2)]$, $0 < x < 1$.

31. Solve the heat equation in spherical coordinates

$$\frac{\partial u}{\partial t} = \frac{a^2}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial u}{\partial r}\right) = \frac{a^2}{r} \frac{\partial^2 (ru)}{\partial r^2}, \qquad 0 \le r < 1, \quad 0 < t,$$

subject to the boundary conditions $\lim_{r\to 0} |u(r,t)| < \infty$, and $u(1,t) = 0$, $0 < t$, and the initial condition $u(r,0) = 1$, $0 \le r < 1$.

32. Solve the heat equation in spherical coordinates

$$\frac{\partial u}{\partial t} = \frac{a^2}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial u}{\partial r}\right) = \frac{a^2}{r} \frac{\partial^2 (ru)}{\partial r^2}, \qquad \alpha < r < \beta, \quad 0 < t,$$

[16] Motivated by problems solved in Gomer, R., 1951: Wall reactions and diffusion in static and flow systems. *J. Chem. Phys.*, **19**, 284–289.

subject to the boundary conditions $u(\alpha, t) = u_r(\beta, t) = 0$, $0 < t$, and the initial condition $u(r, 0) = u_0$, $\alpha < r < \beta$.

33. Solve[17] the heat equation in spherical coordinates

$$\frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} \right) = \frac{a^2}{r}\frac{\partial^2 (ru)}{\partial r^2}, \qquad 0 \le r < b, \quad 0 < t,$$

subject to the boundary conditions

$$\lim_{r\to 0} \frac{\partial u(r,t)}{\partial r} \to 0, \quad \text{and} \quad \frac{\partial u(b,t)}{\partial r} = -\frac{A}{b} u(b,t), \qquad 0 < t,$$

and the initial condition $u(r, 0) = u_0$, $0 \le r < b$.

34. Solve[18] the heat equation in cylindrical coordinates

$$\frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} \right), \qquad 0 \le r < b, \quad 0 < t,$$

subject to the boundary conditions $\lim_{r\to 0} |u(r,t)| < \infty$, and $u(b, t) = u_0$, $0 < t$, and the initial condition $u(r, 0) = 0$, $0 \le r < b$.

35. Solve the heat equation in cylindrical coordinates

$$\frac{\partial u}{\partial t} = \frac{a^2}{r}\frac{\partial}{\partial r}\left(r \frac{\partial u}{\partial r} \right), \qquad 0 \le r < b, \quad 0 < t,$$

subject to the boundary conditions $\lim_{r\to 0} |u(r,t)| < \infty$, and $u(b, t) = \theta$, $0 < t$, and the initial condition $u(r, 0) = 1$, $0 \le r < b$.

36. Solve the heat equation in cylindrical coordinates

$$\frac{\partial u}{\partial t} = \frac{a^2}{r}\frac{\partial}{\partial r}\left(r \frac{\partial u}{\partial r} \right), \qquad 0 \le r < 1, \quad 0 < t,$$

subject to the boundary conditions $\lim_{r\to 0} |u(r,t)| < \infty$, and $u(1, t) = 0$, $0 < t$, and the initial condition

$$u(x, 0) = \begin{cases} A, & 0 \le r < b, \\ B, & b < r < 1. \end{cases}$$

37. The equation[19]

$$\frac{\partial u}{\partial t} = \frac{G}{\rho} + \nu \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} \right), \qquad 0 \le r < b, \quad 0 < t,$$

---

[17] Zhou, H., S. Abanades, G. Flamant, D. Gauthier, and J. Lu, 2002: Simulation of heavy metal vaporization dynamics of a fluidized bed. *Chem. Eng. Sci.*, **57**, 2603–2614. See also Mantell, C., M. Rodriguez, and E. Martinez de la Ossa, 2002: Semi-batch extraction of anthocyanins from red grape pomace in packed beds: Experimental results and process modelling. *Chem. Eng. Sci.*, **57**, 3831–3838.

[18] See Destriau, G., 1946: Propagation des charges électriques sur les pellicules faiblement conductrices "problèm plan." *J. Phys. Radium*, **7**, 43–48.

[19] See Szymanski, P., 1932: Quelques solutions exactes des équations de l'hydrodynamique du fluide visqueux dans le cas d'un tube cylindrique. *J. Math. Pures Appl., Ser. 9*, **11**, 67–107.

governs the velocity $u(r,t)$ of an incompressible fluid of density ρ and kinematic viscosity ν flowing in a long circular pipe of radius b with an imposed, constant pressure gradient $-G$. If the fluid is initially at rest, $u(r,0)=0$, $0\le r<b$, and there is no slip at the wall $u(b,t)=0$, $0<t$, find the velocity at any subsequent time and position.

38. Solve the heat equation in cylindrical coordinates

$$\frac{\partial u}{\partial t}=\frac{a^2}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right),\qquad 0\le r<b,\quad 0<t,$$

subject to the boundary conditions $\lim_{r\to 0}|u(r,t)|<\infty$, and $u_r(b,t)=-h\,u(b,t)$, $0<t$, and the initial condition $u(r,0)=b^2-r^2$, $0\le r<b$.

39. Solve[20] the heat equation in cylindrical coordinates

$$\frac{\partial u}{\partial t}=a^2\left(\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}\right)-\kappa u,\qquad 0\le r<L,\quad 0<t,$$

subject to the boundary conditions $\lim_{r\to 0}|u(r,t)|<\infty$, and $u_r(L,t)=-hu(L,t)$, $0<t$, and the initial condition

$$u(r,0)=\begin{cases}0, & 0\le r<b,\\ T_0, & b<r\le L,\end{cases}$$

where $b<L$, and $0<h,\kappa$.

40. In their study of heat conduction within a thermocouple through which a steady current flows, Reich and Madigan[21] solved the following nonhomogeneous heat conduction problem:

$$\frac{\partial u}{\partial t}-a^2\frac{\partial^2 u}{\partial x^2}=J-P\,\delta(x-b),\qquad 0<x<L,\quad 0<t,\quad 0<b<L,$$

where J represents the Joule heating generated by the steady current and the P term represents the heat loss from Peltier cooling.[22] Find $u(x,t)$ if both ends are kept at zero $[u(0,t)=u(L,t)=0]$ and initially the temperature is zero $[u(x,0)=0]$. The interesting aspect of this problem is the presence of the delta function.

Step 1: Assuming that $u(x,t)$ equals the sum of a steady-state solution $w(x)$ and a transient solution $v(x,t)$, show that the steady-state solution is governed by

$$a^2\frac{d^2w}{dx^2}=P\,\delta(x-b)-J,\qquad w(0)=w(L)=0.$$

Step 2: Show that the steady-state solution is

$$w(x)=\begin{cases}Jx(L-x)/2a^2+Ax, & 0<x<b,\\ Jx(L-x)/2a^2+B(L-x), & b<x<L.\end{cases}$$

---

[20] Mack, W., M. Plöchl, and U. Gamer, 2000: Effects of a temperature cycle on an elastic-plastic shrink fit with solid inclusion. *Chinese J. Mech.*, **16**, 23–30.

[21] Reich, A. D., and J. R. Madigan, 1961: Transient response of a thermocouple circuit under steady currents. *J. Appl. Phys.*, **32**, 294–301.

[22] In 1834 Jean Charles Athanase Peltier (1785–1845) discovered that there is a heating or cooling effect, quite apart from ordinary resistance heating, whenever an electric current flows through the junction between two different metals.

Step 3: The temperature must be continuous at $x = b$; otherwise, we would have infinite heat conduction there. Use this condition to show that $Ab = B(L - b)$.

Step 4: To find a second relationship between A and B, integrate the steady-state differential equation across the interface at $x = b$ and show that

$$\lim_{\epsilon \to 0} a^2 \left. \frac{dw}{dx} \right|_{b-\epsilon}^{b+\epsilon} = P.$$

Step 5: Using the result from Step 4, show that $A + B = -P/a^2$, and

$$w(x) = \begin{cases} Jx(L-x)/2a^2 - Px(L-b)/a^2L, & 0 < x < b, \\ Jx(L-x)/2a^2 - Pb(L-x)/a^2L, & b < x < L. \end{cases}$$

Step 6: Re-express $w(x)$ as a half-range Fourier sine expansion and show that

$$w(x) = \frac{4JL^2}{a^2\pi^3} \sum_{m=1}^{\infty} \frac{\sin[(2m-1)\pi x/L]}{(2m-1)^3} - \frac{2LP}{a^2\pi^2} \sum_{n=1}^{\infty} \frac{\sin(n\pi b/L)\sin(n\pi x/L)}{n^2}.$$

Step 7: Use separation of variables to find the transient solution by solving

$$\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 v}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

subject to the boundary conditions $v(0,t) = v(L,t) = 0$, $0 < t$, and the initial condition $v(x,0) = -w(x), 0 < x < L$.

Step 8: Add the steady-state and transient solutions together and show that

$$u(x,t) = \frac{4JL^2}{a^2\pi^3} \sum_{m=1}^{\infty} \frac{\sin[(2m-1)\pi x/L]}{(2m-1)^3} \left[1 - e^{-a^2(2m-1)^2\pi^2 t/L^2}\right]$$
$$- \frac{2LP}{a^2\pi^2} \sum_{n=1}^{\infty} \frac{\sin(n\pi b/L)\sin(n\pi x/L)}{n^2} \left[1 - e^{-a^2 n^2 \pi^2 t/L^2}\right].$$

41. Use separation of variables to solve[23] the partial differential equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 2a\frac{\partial u}{\partial x}, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions that $u_x(0,t) + 2au(0,t) = 0$, $u_x(1,t) + 2au(1,t) = 0$, $0 < t$, and the initial condition that $u(x,0) = 1$, $0 < x < 1$.

Step 1: Introducing $u(x,t) = e^{-ax}v(x,t)$, show that the problem becomes

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} - a^2 v, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions that $v_x(0,t) + av(0,t) = 0$, $v_x(1,t) + av(1,t) = 0$, $0 < t$, and the initial condition that $u(x,0) = e^{ax}$, $0 < x < 1$.

[23] See DeGroot, S. R., 1942: Théorie phénoménologique de l'effet Soret. *Physica*, **9**, 699–707.

Step 2: Assuming that $v(x,t) = X(x)T(t)$, show that the problem reduces to the ordinary differential equations

$$X'' + (\lambda - a^2)X = 0, \qquad X'(0) + aX(0) = 0, \quad X'(1) + aX(1) = 0,$$

and $T' + \lambda T = 0$, where λ is the separation constant.

Step 3: Solve the eigenvalue problem and show that $\lambda_0 = 0$, $X_0(x) = e^{-ax}$, $T_0(t) = A_0$, and $\lambda_n = a^2 + n^2\pi^2$, $X_n(x) = a\sin(n\pi x) - n\pi\cos(n\pi x)$, and $T_n(t) = A_n e^{-(a^2+n^2\pi^2)t}$, where $n = 1, 2, 3, \ldots$, so that

$$v(x,t) = A_0 e^{-ax} + \sum_{n=1}^{\infty} A_n \left[a\sin(n\pi x) - n\pi\cos(n\pi x)\right] e^{-(a^2+n^2\pi^2)t}.$$

Step 4: Evaluate A_0 and A_n and show that

$$u(x,t) = \frac{2ae^{-2ax}}{1 - e^{-2a}} + 4a\pi \sum_{n=1}^{\infty} \frac{n\left[1 - (-1)^n e^a\right]}{(a^2 + n^2\pi^2)^2} \left[a\sin(n\pi x) - n\pi\cos(n\pi x)\right] e^{-ax-(a^2+n^2\pi^2)t}.$$

42. Use separation of variables to solve[24] the partial differential equation

$$\frac{\partial^3 u}{\partial x^2 \partial t} = \frac{\partial^4 u}{\partial x^4}, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions that $u(0,t) = u_x(0,t) = u(1,t) = u_x(1,t) = 0$, $0 < t$, and the initial condition that $u(x,0) = Ax/2 - (1-A)x^2\left(\frac{3}{2} - x\right)$, $0 \le x \le 1$.

Step 1: Assuming that $u(x,t) = X(x)T(t)$, show that the problem reduces to the ordinary differential equations $X'''' + k^2 X'' = 0$, $X(0) = X'(0) = X(1) = X'(1) = 0$, and $T' + k^2 T = 0$, where k^2 is the separation constant.

Step 2: Solving the eigenvalue problem first, show that

$$X_n(x) = 1 - \cos(k_n x) + \frac{1 - \cos(k_n)}{k_n - \sin(k_n)}[\sin(k_n x) - k_n x],$$

where k_n denotes the nth root of

$$2 - 2\cos(k) - k\sin(k) = \sin(k/2)[\sin(k/2) - (k/2)\cos(k/2)] = 0.$$

Step 3: Using the results from Step 2, show that there are two classes of eigenfunctions: $\kappa_n = 2n\pi$, $X_n(x) = 1 - \cos(2n\pi x)$, and

$$\tan(\kappa_n/2) = \kappa_n/2, \qquad X_n(x) = 1 - \cos(\kappa_n x) + \frac{2}{\kappa_n}[\sin(\kappa_n x) - \kappa_n x].$$

Step 4: Consider the eigenvalue problem

$$X'''' + \lambda X'' = 0, \qquad 0 < x < 1,$$

[24] See Hamza, E. A., 1999: Impulsive squeezing with suction and injection. *J. Appl. Mech.*, **66**, 945–951.

with the boundary conditions $X(0) = X'(0) = X(1) = X'(1) = 0$. Show that the orthogonality condition for this problem is

$$\int_0^1 X_n'(x)X_m'(x)\,dx = 0, \qquad n \neq m,$$

where $X_n(x)$ and $X_m(x)$ are two distinct eigenfunctions of this problem. Then show that we can construct an eigenfunction expansion for an arbitrary function $f(x)$ via

$$f(x) = \sum_{n=1}^{\infty} C_n X_n(x), \qquad \text{provided} \qquad C_n = \frac{\int_0^1 f'(x)X_n'(x)\,dx}{\int_0^1 [X_n'(x)]^2\,dx}$$

and $f'(x)$ exists over the interval $(0, 1)$. Hint: Follow the proof in Section 6.2 and integrate repeatedly by parts to eliminate the higher derivative terms.

Step 5: Show that

$$\int_0^1 [X_n'(x)]^2\,dx = 2n^2\pi^2,$$

if $X_n(x) = 1 - \cos(2n\pi x)$, and

$$\int_0^1 [X_n'(x)]^2\,dx = \kappa_n^2/2,$$

if $X_n(x) = 1 - \cos(\kappa_n x) + 2[\sin(\kappa_n x) - \kappa_n x]/\kappa_n$. Hint: $\sin(\kappa_n) = \kappa_n[1 + \cos(\kappa_n)]/2$.

Step 6: Use the above results to show that

$$u(x,t) = \sum_{n=1}^{\infty} A_n[1 - \cos(2n\pi x)]e^{-4n^2\pi^2 t} + \sum_{n=1}^{\infty} B_n \left\{1 - \cos(\kappa_n x) - \frac{2}{\kappa_n}[\sin(\kappa_n x) - \kappa_n x]\right\} e^{-\kappa_n^2 t},$$

where A_n is the Fourier coefficient corresponding to the eigenfunction $1 - \cos(2n\pi x)$ while B_n is the Fourier coefficient corresponding to the eigenfunction $1 - \cos(\kappa_n x) - 2[\sin(\kappa_n x) - \kappa_n x]/\kappa_n$.

Step 7: Show that $A_n = 0$ and $B_n = 2(1 - A)/\kappa_n^2$, so that

$$u(x,t) = 2(1 - A)\sum_{n=1}^{\infty} \left\{1 - \cos(\kappa_n x) - \frac{2}{\kappa_n}[\sin(\kappa_n x) - \kappa_n x]\right\} \frac{e^{-\kappa_n^2 t}}{\kappa_n^2}.$$

Hint: $\sin(\kappa_n) = \kappa_n[1+\cos(\kappa_n)]/2$, $\sin(\kappa_n) = 2[1-\cos(\kappa_n)]/\kappa_n$, and $\cos(\kappa_n) = (4-\kappa_n^2)/(4+\kappa_n^2)$.

43. Use separation of variables to solve[25] the partial differential equation

$$\frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2}\right), \qquad 0 \le r < 1, \quad 0 < t,$$

[25] See Littlefield, D. L., 1991: Finite conductivity effects on the MHD instabilities in uniformly elongating plastic jets. *Phys. Fluids, Ser. A*, **3**, 1666–1673.

subject to the boundary conditions that

$$\lim_{r\to 0} |u(r,t)| < \infty, \qquad u(1,t) = K, \qquad 0 < t,$$

and the initial condition that

$$u(r,0) = g(r), \qquad 0 < r < 1.$$

Step 1: Setting $u(r,t) = Kr + v(r,t)$, show that the problem becomes

$$\frac{\partial v}{\partial t} = a^2 \left(\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\frac{\partial v}{\partial r} - \frac{v}{r^2} \right), \qquad 0 \le r < 1, \quad 0 < t,$$

subject to the boundary conditions that

$$\lim_{r\to 0} |v(r,t)| < \infty, \qquad v(1,t) = 0, \qquad 0 < t,$$

and the initial condition that

$$u(r,0) = g(r) - Kr, \qquad 0 < r < 1.$$

Step 2: Assuming that $v(r,t) = R(r)T(t)$, show that the problem reduces to the ordinary differential equations

$$R'' + \frac{1}{r}R' + \left(k^2 - \frac{1}{r^2}\right) R = 0,$$

and

$$T' + a^2k^2T = 0$$

with the boundary conditions $\lim_{r\to 0} |R(r)| < \infty$ and $R(1) = 0$, where k^2 is the separation constant.

Step 3: Solving the eigenvalue problem first, show that $R_n(r) = J_1(k_n r)$ where k_n denotes the nth root of $J_1(k) = 0$ and $n = 1, 2, 3, \ldots$.

Step 4: Show that $T_n(t) = A_n e^{-a^2k_n^2t}$ so that

$$v(r,t) = \sum_{n=1}^{\infty} A_n J_1(k_n r) e^{-a^2k_n^2t}.$$

Step 5: Using the initial condition, show that

$$v(r,0) = g(r) - Kr = \sum_{n=1}^{\infty} A_n J_1(k_n r).$$

Step 6: Evaluate A_n and show that it equals

$$A_n = \frac{2}{k_n J_0(k_n)} \left[K + \frac{k_n}{J_0(k_n)} \int_0^1 g(r) J_1(k_n r)\, r\, dr \right].$$

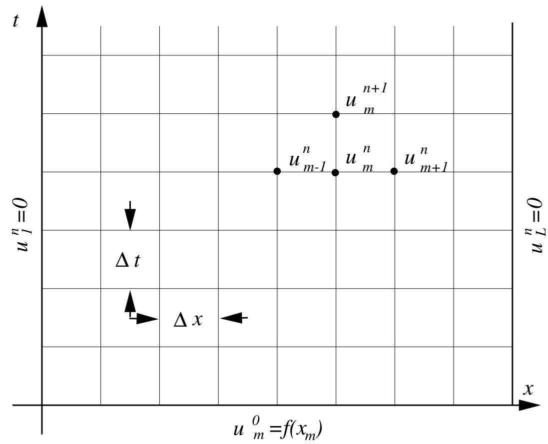


Figure 8.4.1: Schematic of the numerical solution of the heat equation when we hold both ends at a temperature of zero.

8.4 NUMERICAL SOLUTION OF THE HEAT EQUATION

In the previous chapter we showed how we may use finite difference techniques to solve the wave equation. In this section we show that similar considerations hold for the heat equation.

Starting with the heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{8.4.1}$$

we must first replace the exact derivatives with finite differences. Drawing upon our work in Section 7.6,

$$\frac{\partial u(x_m, t_n)}{\partial t} = \frac{u_m^{n+1} - u_m^n}{\Delta t} + O(\Delta t), \tag{8.4.2}$$

and

$$\frac{\partial^2 u(x_m, t_n)}{\partial x^2} = \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{(\Delta x)^2} + O[(\Delta x)^2], \tag{8.4.3}$$

where the notation u_m^n denotes $u(x_m, t_n)$. Figure 8.4.1 illustrates our numerical scheme when we hold both ends at the temperature of zero. Substituting Equation 8.4.2 and Equation 8.4.3 into Equation 8.4.1 and rearranging,

$$u_m^{n+1} = u_m^n + \frac{a^2 \Delta t}{(\Delta x)^2}\left(u_{m+1}^n - 2u_m^n + u_{m-1}^n\right). \tag{8.4.4}$$

The numerical integration begins with $n = 0$ and the value of u_{m+1}^0, u_m^0, and u_{m-1}^0 are given by $f(m\Delta x)$.

Once again we must check the *convergence*, *stability*, and *consistency* of our scheme. We begin by writing u_{m+1}^n, u_{m-1}^n, and u_m^{n+1} in terms of the exact solution u and its derivatives evaluated at the point $x_m = m\Delta x$ and $t_n = n\Delta t$. By Taylor's expansion,

$$u_{m+1}^n = u_m^n + \Delta x \left.\frac{\partial u}{\partial x}\right|_n^m + \tfrac{1}{2}(\Delta x)^2 \left.\frac{\partial^2 u}{\partial x^2}\right|_n^m + \tfrac{1}{6}(\Delta x)^3 \left.\frac{\partial^3 u}{\partial x^3}\right|_n^m + \cdots, \tag{8.4.5}$$

$$u_{m-1}^n = u_m^n - \Delta x \frac{\partial u}{\partial x}\bigg|_n^m + \tfrac{1}{2}(\Delta x)^2 \frac{\partial^2 u}{\partial x^2}\bigg|_n^m - \tfrac{1}{6}(\Delta x)^3 \frac{\partial^3 u}{\partial x^3}\bigg|_n^m + \cdots, \tag{8.4.6}$$

and

$$u_m^{n+1} = u_m^n + \Delta t \frac{\partial u}{\partial t}\bigg|_n^m + \tfrac{1}{2}(\Delta t)^2 \frac{\partial^2 u}{\partial t^2}\bigg|_n^m + \tfrac{1}{6}(\Delta t)^3 \frac{\partial^3 u}{\partial t^3}\bigg|_n^m + \cdots. \tag{8.4.7}$$

Substituting into Equation 8.4.4, we obtain

$$\begin{aligned}\frac{u_m^{n+1} - u_m^n}{\Delta t} &- a^2 \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{(\Delta x)^2} \\ &= \left(\frac{\partial u}{\partial t} - a^2 \frac{\partial^2 u}{\partial x^2} \right)\bigg|_n^m + \tfrac{1}{2}\Delta t \frac{\partial^2 u}{\partial t^2}\bigg|_n^m - \tfrac{1}{12}(a\Delta x)^2 \frac{\partial^4 u}{\partial x^4}\bigg|_n^m + \cdots. \end{aligned} \tag{8.4.8}$$

The first term on the right side of Equation 8.4.8 vanishes because $u(x,t)$ satisfies the heat equation. Thus, in the limit of $\Delta x \to 0$, $\Delta t \to 0$, the right side of Equation 8.4.8 vanishes and the scheme is *consistent*.

To determine the *stability* of the explicit scheme, we again use the Fourier method. Assuming a solution of the form:

$$u_n^m = e^{im\theta} e^{in\lambda}, \tag{8.4.9}$$

we substitute Equation 8.4.9 into Equation 8.4.4 and find that

$$\frac{e^{i\lambda} - 1}{\Delta t} = a^2 \frac{e^{i\theta} - 2 + e^{-i\theta}}{(\Delta x)^2}, \tag{8.4.10}$$

or

$$e^{i\lambda} = 1 - 4 \frac{a^2 \Delta t}{(\Delta x)^2} \sin^2\left(\frac{\theta}{2}\right). \tag{8.4.11}$$

The quantity $e^{i\lambda}$ will grow exponentially unless

$$-1 \le 1 - 4 \frac{a^2 \Delta t}{(\Delta x)^2} \sin^2\left(\frac{\theta}{2}\right) < 1. \tag{8.4.12}$$

The right inequality is trivially satisfied if $a^2\Delta t/(\Delta x)^2 > 0$, while the left inequality yields

$$\frac{a^2 \Delta t}{(\Delta x)^2} \le \frac{1}{2\sin^2(\theta/2)}, \tag{8.4.13}$$

leading to the stability condition $0 < a^2\Delta t/(\Delta x)^2 \le \frac{1}{2}$. This is a rather restrictive condition because doubling the resolution (halving Δx) requires that we reduce the time step by a quarter. Thus, for many calculations the required time step may be unacceptably small. For this reason, many use an implicit form of the finite differencing (Crank-Nicholson implicit method[26]):

$$\frac{u_m^{n+1} - u_m^n}{\Delta t} = \frac{a^2}{2} \left[\frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{(\Delta x)^2} + \frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{(\Delta x)^2} \right], \tag{8.4.14}$$

[26] Crank, J., and P. Nicholson, 1947: A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Proc. Cambridge. Philos. Soc.*, **43**, 50–67.

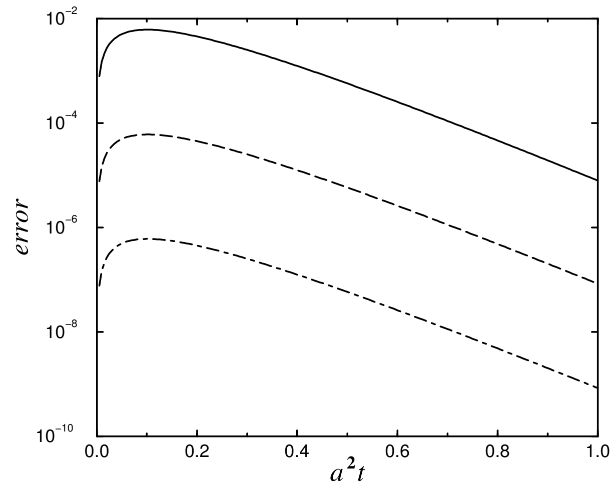


Figure 8.4.2: The growth of error $||e_n||$ as a function of a^2t for various resolutions. For the top line, $\Delta x = 0.1$; for the middle line, $\Delta x = 0.01$; and for the bottom line, $\Delta x = 0.001$.

although it requires the solution of a simultaneous set of linear equations. However, there are several efficient methods for their solution.

Finally we must check and see if our explicit scheme *converges* to the true solution. If we let e_m^n denote the difference between the exact and our finite differenced solution to the heat equation, we can use Equation 8.4.8 to derive the equation governing e_m^n and find that

$$e_m^{n+1} = e_m^n + \frac{a^2\Delta t}{(\Delta x)^2}\left(e_{m+1}^n - 2e_m^n + e_{m-1}^n\right) + O[(\Delta t)^2 + \Delta t(\Delta x)^2], \tag{8.4.15}$$

for $m = 1, 2, \ldots, M$. Assuming that $a^2\Delta t/(\Delta x)^2 \leq \frac{1}{2}$, then

$$\left|e_m^{n+1}\right| \leq \frac{a^2\Delta t}{(\Delta x)^2}\left|e_{m-1}^n\right| + \left[1 - 2\frac{a^2\Delta t}{(\Delta x)^2}\right]\left|e_m^n\right| + \frac{a^2\Delta t}{(\Delta x)^2}\left|e_{m+1}^n\right| + A[(\Delta t)^2 + \Delta t(\Delta x)^2] \tag{8.4.16}$$

$$\leq ||e_n|| + A[(\Delta t)^2 + \Delta t(\Delta x)^2], \tag{8.4.17}$$

where $||e_n|| = \max_{m=0,1,\ldots,M}\left|e_m^n\right|$. Consequently,

$$||e_{n+1}|| \leq ||e_n|| + A[(\Delta t)^2 + \Delta t(\Delta x)^2]. \tag{8.4.18}$$

Because $||e_0|| = 0$ and $n\Delta t \leq t_n$, we find that

$$||e_{n+1}|| \leq An[(\Delta t)^2 + \Delta t(\Delta x)^2] \leq At_n[\Delta t + (\Delta x)^2]. \tag{8.4.19}$$

As $\Delta x \to 0$, $\Delta t \to 0$, the errors tend to zero and we have convergence. We have illustrated Equation 8.4.19 in Figure 8.4.2 by using the finite difference equation, Equation 8.4.4, to compute $||e_n||$ during a numerical experiment that used $a^2\Delta t/(\Delta x)^2 = 0.5$, and $f(x) = \sin(\pi x)$. Note how each increase of resolution by 10 results in a drop in the error by 100.

The following examples illustrate the use of numerical methods.

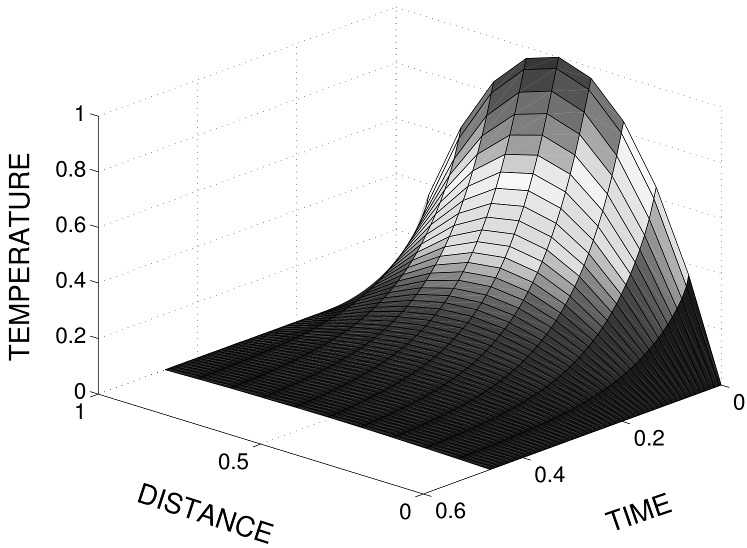


Figure 8.4.3: The numerical solution $u(x,t)$ of the heat equation with $a^2\Delta t/(\Delta x)^2 = 0.47$ at various positions $x' = x/L$ and times $t' = a^2t/L^2$ using Equation 8.4.4. The initial temperature $u(x,0)$ equals $4x'(1-x')$ and we hold both ends at a temperature of zero.

• **Example 8.4.1**

For our first example, we redo Example 8.3.1 with $a^2\Delta t/(\Delta x)^2 = 0.47$ and 0.53. Our numerical solution was computed using the MATLAB script

```
clear
coeff = 0.47; % coeff = a^2Δt/(Δx)^2
ncount = 1; dx = 0.1; dt = coeff * dx * dx;
N = 99; x = 0:dx:1;
M = 1/dx + 1; % M = number of spatial grid points
tplot(1) = 0; u = zeros(M,N+1);
for m = 1:M; u(m,1)=4*x(m)*(1-x(m)); temp(m,1)=u(m,1); end
% integrate forward in time
for n = 1:N
  t = dt * n;
  for m = 2:M-1
    u(m,n+1) = u(m,n) + coeff*(u(m+1,n)-2*u(m,n)+u(m-1,n));
  end
  if mod(n+1,2) == 0
    ncount = ncount + 1; tplot(ncount) = t;
    for m = 1:M; temp(m,ncount) = u(m,n+1); end
end; end
% plot the numerical solution
X = x' * ones(1,length(tplot)); T = ones(M,1) * tplot;
surf(X,T,temp)
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('TEMPERATURE','Fontsize',20)
```

As Figure 8.4.3 shows, the solution with $a^2\Delta t/(\Delta x)^2 < 1/2$ performs well. On the other hand, Figure 8.4.4 shows small-scale, growing disturbances when $a^2\Delta t/(\Delta x)^2 > 1/2$. It

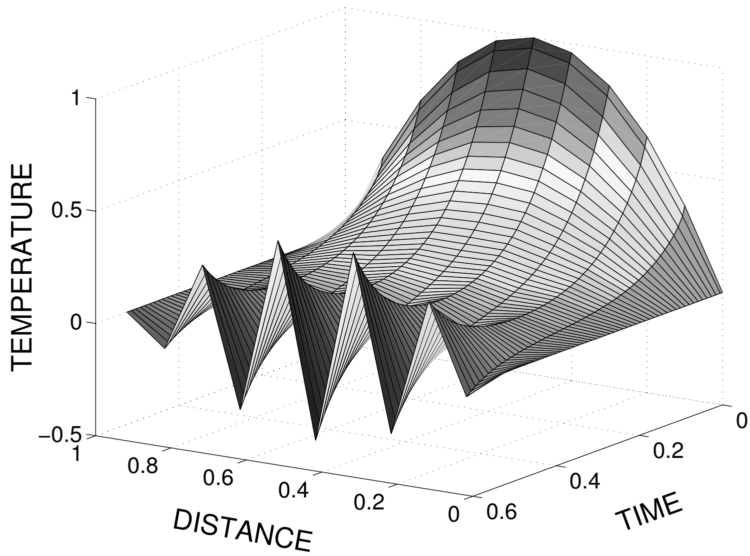


Figure 8.4.4: Same as Figure 8.4.3 except that $a^2\Delta t/(\Delta x)^2 = 0.53$.

should be noted that for the reasonable $\Delta x = L/100$, it takes approximately *20,000* time steps before we reach $a^2t/L^2 = 1$. □

• Example 8.4.2

In this example, we redo the previous example with an insulated end at $x = L$. Using the centered differencing formula,

$$u_{M+1}^n - u_{M-1}^n = 0, \tag{8.4.20}$$

because $u_x(L,t) = 0$. Also, at $i = M$,

$$u_M^{n+1} = u_M^n + \frac{a^2\Delta t}{(\Delta x)^2}\left(u_{M+1}^n - 2u_M^n + u_{M-1}^n\right). \tag{8.4.21}$$

Eliminating u_{M+1}^n between the two equations,

$$u_M^{n+1} = u_M^n + \frac{a^2\Delta t}{(\Delta x)^2}\left(2u_{M-1}^n - 2u_M^n\right). \tag{8.4.22}$$

To implement this new boundary condition in our MATLAB script, we add the line

```
u(M,n+1) = u(M,n) + 2 * coeff * (u(M-1,n) - u(M,n));
```

after the lines

```
for m = 2:M-1
u(m,n+1) = u(m,n) + coeff * (u(m+1,n) - 2 * u(m,n) + u(m-1,n));
end
```

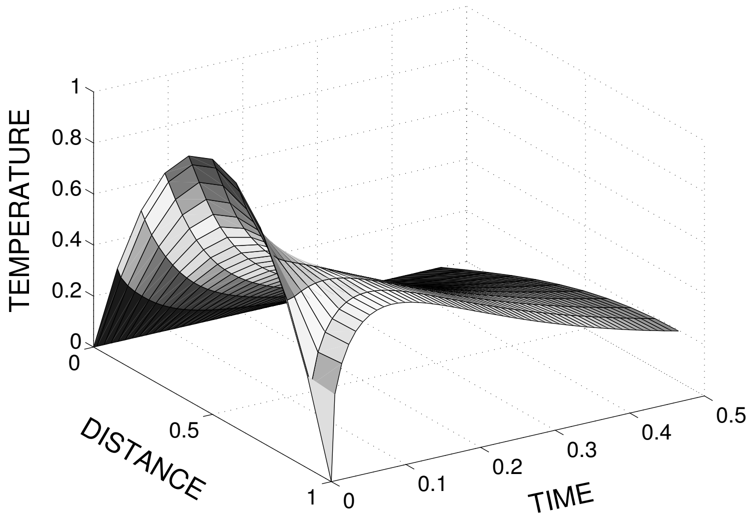


Figure 8.4.5: Same as Figure 8.4.4 except that we now have an insulated boundary condition $u_x(L, t) = 0$.

Figure 8.4.5 illustrates our numerical solution at various positions and times.

Project: Implicit Numerical Integration of the Heat Equation

The difficulty in using explicit time differencing to solve the heat equation is the very small time step that must be taken at moderate spatial resolutions to ensure stability. This small time step translates into an unacceptably long execution time. In this project you will investigate the Crank-Nicholson implicit scheme, which allows for a much more reasonable time step.

Step 1: Develop a MATLAB script that uses the Crank-Nicholson equation, Equation 8.4.14, to numerically integrate the heat equation. To do this, you will need a tridiagonal solver to find u_m^{n+1}. This is explained at the end of Section 3.1. However, many numerical methods books[27] actually have code already developed for your use. You might as well use this code.

Step 2: Test your code by solving the heat equation given the initial condition $u(x, 0) = \sin(\pi x)$, and the boundary conditions $u(0, t) = u(1, t) = 0$. Find the solution for various values of Δt with $\Delta x = 0.01$. Compare this numerical solution against the exact solution that you can find. How does the error (between the numerical and exact solutions) change with Δt? For small Δt, the errors should be small. If not, then you have a mistake in your code.

Step 3: Once you have confidence in your code, discuss the behavior of the scheme for various values of Δx and Δt for the initial condition $u(x, 0) = 0$ for $0 \le x < \frac{1}{2}$, and $u(x, 0) = 1$ for $\frac{1}{2} < x \le 1$ with the boundary conditions $u(0, t) = u(1, t) = 0$. See Figure 8.4.6. Although you can take quite a large Δt, what happens? Did a similar problem arise

[27] For example, Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1986: *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Section 2.6.

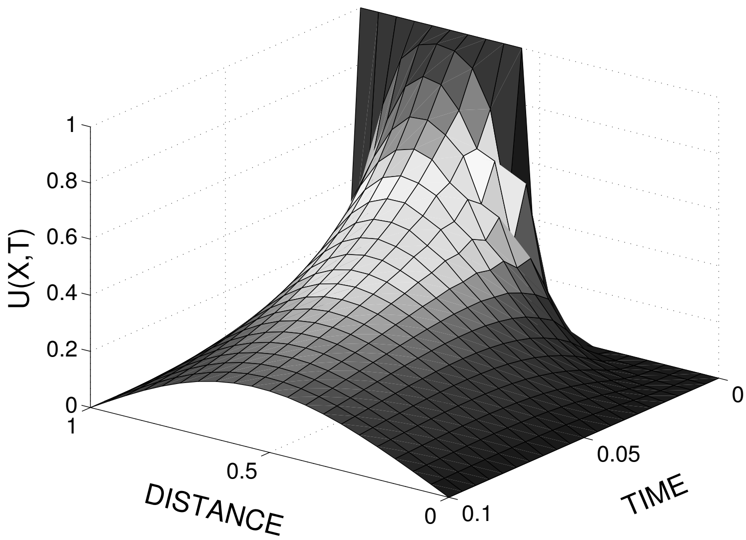


Figure 8.4.6: The numerical solution $u(x,t)$ of the heat equation $u_t = a^2 u_{xx}$ using the Crank-Nicholson method. The parameters used in the numerical solution are $a^2\Delta t = 0.005$ and $\Delta x = 0.05$. Both ends are held at zero with an initial condition of $u(x,0) = 0$ for $0 \le x < \frac{1}{2}$, and $u(x,0) = 1$ for $\frac{1}{2} < x \le 1$.

in Step 2? Explain your results.[28] Zvan et al.[29] have reported a similar problem in the numerical integration of the Black-Scholes equation from mathematical finance.

Further Readings

Carslaw, H. S., and J. C. Jaeger, 1959: *Conduction of Heat in Solids.* Oxford University Press, 510 pp. The source book on solving the heat equation.

Crank, J., 1970: *The Mathematics of Diffusion.* Oxford University Press, 347 pp. A source book on the solution of the heat equation.

Koshlyakov, N. S., M. M. Smirnov, and E. B. Gliner, 1964: *Differential Equations of Mathematical Physics.* North-Holland Publishing, 701 pp. See Part III. Nice presentation of mathematical techniques.

Morse, P. M., and H. Feshback, 1953: *Methods of Theoretical Physics.* McGraw-Hill Book Co., 997 pp. A portion of Chapter 12 is devoted to solving the heat equation.

[28] Luskin, M., and R. Rannacher, 1982: On the smoothing property of the Crank-Nicolson scheme. *Applicable Anal.*, **14**, 117–135.

[29] Zvan, R., K. Vetzal, and P. Forsyth, 1998: Swing low, swing high. *Risk*, **11(3)**, 71–75.

Chapter 9

Laplace's Equation

In the previous chapter we solved the one-dimensional heat equation. Quite often we found that the transient solution died away, leaving a steady state. The partial differential equation that describes the steady state for two-dimensional heat conduction is Laplace's equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \qquad (\mathbf{9.0.1})$$

In general, this equation governs physical processes where *equilibrium* has been reached. It also serves as the prototype for a wider class of *elliptic equations*

$$a(x,t)\frac{\partial^2 u}{\partial x^2} + b(x,t)\frac{\partial^2 u}{\partial x \partial t} + c(x,t)\frac{\partial^2 u}{\partial t^2} = f\left(x, t, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial t}\right), \qquad (\mathbf{9.0.2})$$

where $b^2 < 4ac$. Unlike the heat and wave equations, there are no initial conditions and the boundary conditions completely specify the solution. In this chapter we present some of the common techniques for solving this equation.

9.1 DERIVATION OF LAPLACE'S EQUATION

Imagine a thin, flat plate of heat-conducting material between two sheets of insulation. Sufficient time has passed so that the temperature depends only on the spatial coordinates x and y. Let us now apply the law of conservation of energy (in rate form) to a small rectangle with sides Δx and Δy.

If $q_x(x, y)$ and $q_y(x, y)$ denote the heat flow rates in the x- and y-direction, respectively, conservation of energy requires that the heat flow into the slab equals the heat flow out of the slab if there is no storage or generation of heat. Now

$$\text{rate in} = q_x(x, y + \Delta y/2)\Delta y + q_y(x + \Delta x/2, y)\Delta x, \qquad (\mathbf{9.1.1})$$

and

$$\text{rate out} = q_x(x + \Delta x, y + \Delta y/2)\Delta y + q_y(x + \Delta x/2, y + \Delta y)\Delta x. \qquad (\mathbf{9.1.2})$$

If the plate has unit thickness,

$$[q_x(x, y + \Delta y/2) - q_x(x + \Delta x, y + \Delta y/2)]\Delta y + [q_y(x + \Delta x/2, y) - q_y(x + \Delta x/2, y + \Delta y)]\Delta x = 0. \qquad (\mathbf{9.1.3})$$

Upon dividing through by $\Delta x \Delta y$, we obtain two differences quotients on the left side of Equation 9.1.3. In the limit as $\Delta x, \Delta y \to 0$, they become partial derivatives, giving

$$\frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = 0 \qquad (\mathbf{9.1.4})$$

for any point (x, y).

We now employ Fourier's law to eliminate the rates q_x and q_y, yielding

$$\frac{\partial}{\partial x}\left(a^2 \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(a^2 \frac{\partial u}{\partial y}\right) = 0, \qquad (\mathbf{9.1.5})$$

if we have an isotropic (same in all directions) material. Finally, if a^2 is constant, Equation 9.1.5 reduces to

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \qquad (\mathbf{9.1.6})$$

which is the two-dimensional, steady-state heat equation (i.e., $u_t \approx 0$ as $t \to \infty$).

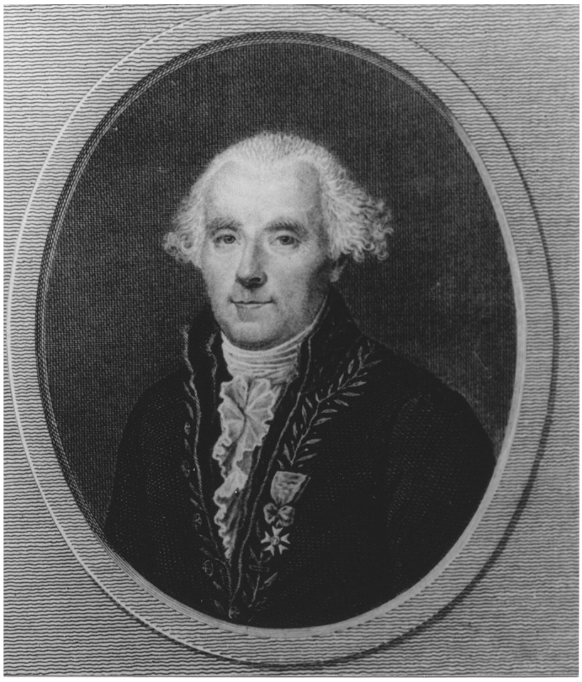
Solutions of Laplace's equation (called *harmonic functions*) differ fundamentally from those encountered with the heat and wave equations. These latter two equations describe the evolution of some phenomena. Laplace's equation, on the other hand, describes things at equilibrium. Consequently, any change in the boundary conditions affects to some degree the *entire* domain because a change to any one point causes its neighbors to change in order to reestablish the equilibrium. Those points will, in turn, affect others. Because all of these points are in equilibrium, this modification must occur instantaneously.

Further insight follows from the *maximum principle.* If Laplace's equation governs a region, then its solution cannot have a relative maximum or minimum *inside* the region unless the solution is constant.[1] If we think of the solution as a steady-state temperature distribution, this principle is clearly true because at any one point the temperature cannot be greater than at all other nearby points. If that were so, heat would flow away from the hot point to cooler points nearby, thus eliminating the hot spot when equilibrium was once again restored.

It is often useful to consider the two-dimensional Laplace's equation in other coordinate systems. In polar coordinates, where $x = r\cos(\theta)$, $y = r\sin(\theta)$, and $z = z$, Laplace's equation becomes

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \qquad (\mathbf{9.1.7})$$

[1] For the proof, see Courant, R., and D. Hilbert, 1962: *Methods of Mathematical Physics, Vol. 2: Partial Differential Equations*. Interscience, pp. 326–331.



Today we best remember Pierre-Simon Laplace (1749–1827) for his work in celestial mechanics and probability. In his five volumes *Traité de Mécanique céleste* (1799–1825), he accounted for the theoretical orbits of the planets and their satellites. Laplace's equation arose during this study of gravitational attraction. (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

if the problem possesses axisymmetry. On the other hand, if the solution is independent of z, Laplace's equation becomes

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0. \tag{9.1.8}$$

In spherical coordinates, $x = r\cos(\varphi)\sin(\theta)$, $y = r\sin(\varphi)\sin(\theta)$, and $z = r\cos(\theta)$, where $r^2 = x^2 + y^2 + z^2$, θ is the angle measured *down* to the point from the z-axis (colatitude) and φ is the angle made between the x-axis and the projection of the point on the xy plane. In the case of axisymmetry (no φ dependence), Laplace's equation becomes

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}\left[\sin(\theta)\frac{\partial u}{\partial \theta}\right] = 0. \tag{9.1.9}$$

9.2 BOUNDARY CONDITIONS

Because Laplace's equation involves time-independent phenomena, we must only specify boundary conditions. As we discussed in Section 8.2, we can classify these boundary conditions as follows:

1. Dirichlet condition: u given

2. Neumann condition: $\dfrac{\partial u}{\partial n}$ given, where n is the unit normal direction

3. Robin condition: $u + \alpha \dfrac{\partial u}{\partial n}$ given

along any section of the boundary. In the case of Laplace's equation, if all of the boundaries have Neumann conditions, then the solution is not unique. This follows from the fact that if $u(x, y)$ is a solution, so is $u(x, y) + c$, where c is any constant.

Finally we note that we must specify the boundary conditions along each side of the boundary. These sides may be at infinity as in problems with semi-infinite domains. We must specify values along the entire boundary because we could not have an equilibrium solution if any portion of the domain was undetermined.

9.3 SEPARATION OF VARIABLES

As in the case of the heat and wave equations, separation of variables is the most popular technique for solving Laplace's equation. Although the same general procedure carries over from the previous two chapters, the following examples fill out the details.

• Example 9.3.1: Groundwater flow in a valley

Over a century ago, a French hydraulic engineer named Henri-Philibert-Gaspard Darcy (1803–1858) published the results of a laboratory experiment on the flow of water through sand. He showed that the *apparent* fluid velocity $\mathbf{q}$ relative to the sand grains is directly proportional to the gradient of the hydraulic potential $-k\nabla\varphi$, where the hydraulic potential φ equals the sum of the elevation of the point of measurement plus the pressure potential $(p/\rho g)$. In the case of steady flow, the combination of Darcy's law with conservation of mass $\nabla \cdot \mathbf{q} = 0$ yields Laplace's equation $\nabla^2\varphi = 0$ if the aquifer is isotropic (same in all directions) and homogeneous.

To illustrate how separation of variables can be used to solve Laplace's equation, we will determine the hydraulic potential within a small drainage basin that lies in a shallow valley. See Figure 9.3.1. Following Tóth,[2] the governing equation is the two-dimensional Laplace equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < L, \quad 0 < y < z_0, \tag{9.3.1}$$

along with the boundary conditions

$$u(x, z_0) = gz_0 + gcx, \tag{9.3.2}$$

$$u_x(0, y) = u_x(L, y) = 0, \qquad \text{and} \qquad u_y(x, 0) = 0, \tag{9.3.3}$$

where $u(x, y)$ is the hydraulic potential, g is the acceleration due to gravity, and c gives the slope of the topography. The conditions $u_x(L, y) = 0$, and $u_y(x, 0) = 0$ specify a no-flow condition through the bottom and sides of the aquifer. The condition $u_x(0, y) = 0$ ensures symmetry about the $x = 0$ line. Equation 9.3.1 gives the fluid potential at the water table, where z_0 is the elevation of the water table above the standard datum. The term gcx in

---

[2] Tóth, J., 1962: A theory of groundwater motion in small drainage basins in central Alberta, Canada. *J. Geophys. Res.*, **67**, 4375–4387.

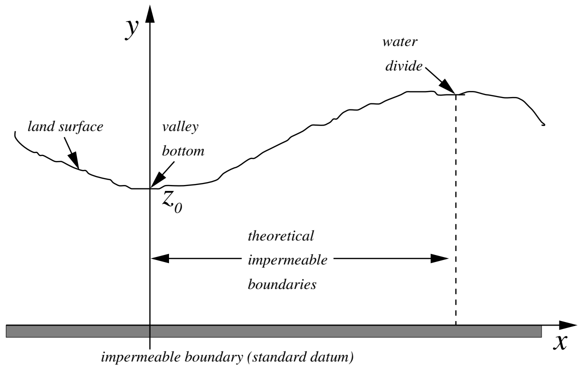


Figure 9.3.1: Cross section of a valley.

Equation 9.3.2 expresses the increase of the potential from the valley bottom toward the water divide. On average it closely follows the topography.

Following the pattern set in the previous two chapters, we assume that $u(x, y) = X(x)Y(y)$. Then Equation 9.3.1 becomes

$$X''Y + XY'' = 0. \tag{9.3.4}$$

Separating the variables yields

$$\frac{X''}{X} = -\frac{Y''}{Y}. \tag{9.3.5}$$

Both sides of Equation 9.3.5 must be constant, but the sign of that constant is not obvious. From previous experience we anticipate that the ordinary differential equation in the x-direction leads to a Sturm-Liouville problem because it possesses homogeneous boundary conditions. Proceeding along this line of reasoning, we consider three separation constants.

Trying a positive constant (say, m^2), Equation 9.3.5 separates into the two ordinary differential equations

$$X'' - m^2X = 0, \qquad \text{and} \qquad Y'' + m^2Y = 0, \tag{9.3.6}$$

which have the solutions

$$X(x) = A\cosh(mx) + B\sinh(mx), \tag{9.3.7}$$

and

$$Y(y) = C\cos(my) + D\sin(my). \tag{9.3.8}$$

Because the boundary conditions, Equation 9.3.3, imply $X'(0) = X'(L) = 0$, both A and B must be zero, leading to the trivial solution $u(x, y) = 0$.

When the separation constant equals zero, we find a nontrivial solution given by the eigenfunction $X_0(x) = 1$, and $Y_0(y) = \frac{1}{2}A_0 + B_0y$. However, because $Y_0'(0) = 0$ from Equation 9.3.3, $B_0 = 0$. Thus, the particular solution for a zero separation constant is $u_0(x, y) = A_0/2$.

Finally, taking both sides of Equation 9.3.5 equal to $-k^2$,

$$X'' + k^2X = 0, \qquad \text{and} \qquad Y'' - k^2Y = 0. \tag{9.3.9}$$

The first of these equations, along with the boundary conditions $X'(0) = X'(L) = 0$, gives the eigenfunction $X_n(x) = \cos(k_n x)$, with $k_n = n\pi/L$, $n = 1, 2, 3, \ldots$. The function $Y_n(y)$ for the same separation constant is

$$Y_n(y) = A_n \cosh(k_n y) + B_n \sinh(k_n y). \tag{9.3.10}$$

We must take $B_n = 0$ because $Y_n'(0) = 0$.

We now have the product solution $X_n(x)Y_n(y)$, which satisfies Laplace's equation and all of the boundary conditions except Equation 9.3.2. By the principle of superposition, the general solution is

$$u(x, y) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi x}{L}\right) \cosh\left(\frac{n\pi y}{L}\right). \tag{9.3.11}$$

Applying Equation 9.3.2, we find that

$$u(x, z_0) = gz_0 + gcx = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi x}{L}\right) \cosh\left(\frac{n\pi z_0}{L}\right), \tag{9.3.12}$$

which we recognize as a Fourier half-range cosine series such that

$$A_0 = \frac{2}{L} \int_0^L (gz_0 + gcx)\, dx, \tag{9.3.13}$$

and

$$\cosh\left(\frac{n\pi z_0}{L}\right) A_n = \frac{2}{L} \int_0^L (gz_0 + gcx) \cos\left(\frac{n\pi x}{L}\right)\, dx. \tag{9.3.14}$$

Performing the integrations,

$$A_0 = 2gz_0 + gcL, \tag{9.3.15}$$

and

$$A_n = -\frac{2gcL[1 - (-1)^n]}{n^2\pi^2 \cosh(n\pi z_0/L)}. \tag{9.3.16}$$

Finally, the complete solution is

$$u(x, y) = gz_0 + \frac{gcL}{2} - \frac{4gcL}{\pi^2} \sum_{m=1}^{\infty} \frac{\cos[(2m-1)\pi x/L] \cosh[(2m-1)\pi y/L]}{(2m-1)^2 \cosh[(2m-1)\pi z_0/L]}. \tag{9.3.17}$$

Figure 9.3.2 presents two graphs by Tóth for two different aquifers. We see that the solution satisfies the boundary condition at the bottom and side boundaries. Water flows from the elevated land (on the right) into the valley (on the left), from regions of high to low hydraulic potential. □

• Example 9.3.2

In the previous example, we had the advantage of homogeneous boundary conditions along $x = 0$ and $x = L$. In a different hydraulic problem, Kirkham[3] solved the more difficult problem of

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < L, \quad 0 < y < h, \tag{9.3.18}$$

[3] Kirkham, D., 1958: Seepage of steady rainfall through soil into drains. *Trans. Am. Geophys. Union*, **39**, 892–908.

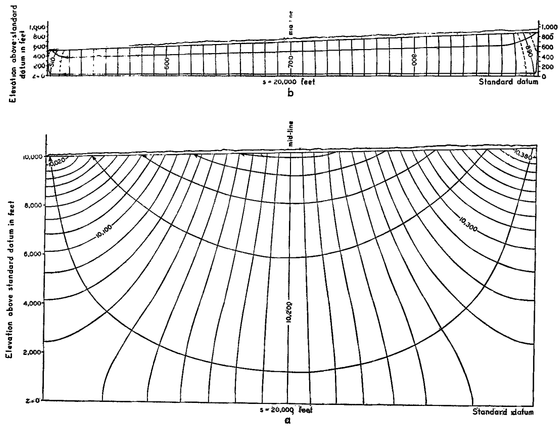


Figure 9.3.2: Two-dimensional potential distribution and flow patterns for different depths of the horizontally impermeable boundary.

subject to the Dirichlet boundary conditions

$$u(x,0) = Rx, \quad u(x,h) = RL, \quad u(L,y) = RL, \tag{9.3.19}$$

and

$$u(0,y) = \begin{cases} 0, & 0 < y < a, \\ \frac{RL}{b-a}(y-a), & a < y < b, \\ RL, & b < y < h. \end{cases} \tag{9.3.20}$$

This problem arises in finding the steady flow within an aquifer resulting from the introduction of water at the top due to a steady rainfall and its removal along the sides by drains. The parameter L equals half of the distance between the drains, h is the depth of the aquifer, and R is the rate of rainfall.

The point of this example is: *We need homogeneous boundary conditions along either the x or y boundaries for separation of variables to work.* We achieve this by breaking the original problem into two parts, namely

$$u(x,y) = v(x,y) + w(x,y) + RL, \tag{9.3.21}$$

where

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0, \quad 0 < x < L, \quad 0 < y < h, \tag{9.3.22}$$

with

$$v(0,y) = v(L,y) = 0, \qquad v(x,h) = 0, \tag{9.3.23}$$

and

$$v(x,0) = R(x-L); \tag{9.3.24}$$

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = 0, \quad 0 < x < L, \quad 0 < y < h, \tag{9.3.25}$$

with

$$w(x,0) = w(x,h) = 0, \qquad w(L,y) = 0, \tag{9.3.26}$$

and

$$w(0,y) = \begin{cases} -RL, & 0 < y < a, \\ \frac{RL}{b-a}(y-a) - RL, & a < y < b, \\ 0, & b < y < h. \end{cases} \tag{9.3.27}$$

Employing the same technique as in Example 9.3.1, we find that

$$v(x,y) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right) \frac{\sinh[n\pi(h-y)/L]}{\sinh(n\pi h/L)}, \tag{9.3.28}$$

where

$$A_n = \frac{2}{L}\int_0^L R(x-L)\sin\left(\frac{n\pi x}{L}\right)\,dx = -\frac{2RL}{n\pi}. \tag{9.3.29}$$

Similarly, the solution to $w(x,y)$ is found to be

$$w(x,y) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi y}{h}\right) \frac{\sinh[n\pi(L-x)/h]}{\sinh(n\pi L/h)}, \tag{9.3.30}$$

where

$$B_n = \frac{2}{h}\left[-RL\int_0^a \sin\left(\frac{n\pi y}{h}\right)\,dy + RL\int_a^b \left(\frac{y-a}{b-a}-1\right)\sin\left(\frac{n\pi y}{h}\right)\,dy\right] \tag{9.3.31}$$

$$= \frac{2RL}{\pi}\left\{\frac{h}{(b-a)n^2\pi}\left[\sin\left(\frac{n\pi b}{h}\right) - \sin\left(\frac{n\pi a}{h}\right)\right] - \frac{1}{n}\right\}. \tag{9.3.32}$$

The complete solution consists of substituting Equation 9.3.28 and Equation 9.3.30 into Equation 9.3.21. □

• Example 9.3.3

The *electrostatic potential* is defined as the amount of work that must be done against electric forces to bring a unit charge from a reference point to a given point. It is readily shown[4] that the electrostatic potential is described by Laplace's equation if there is no charge within the domain. Let us find the electrostatic potential $u(r,z)$ inside a closed cylinder of length L and radius a. The base and lateral surfaces have the potential 0 while the upper surface has the potential V.

Because the potential varies in only r and z, Laplace's equation in cylindrical coordinates reduces to

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < a, \quad 0 < z < L, \tag{9.3.33}$$

subject to the boundary conditions

$$u(a,z) = u(r,0) = 0, \quad \text{and} \quad u(r,L) = V. \tag{9.3.34}$$

[4] For static fields, $\nabla \times \mathbf{E} = \mathbf{0}$, where $\mathbf{E}$ is the electric force. From Section 4.4, we can introduce a potential φ such that $\mathbf{E} = \nabla\varphi$. From Gauss' law, $\nabla \cdot \mathbf{E} = \nabla^2\varphi = 0$.

To solve this problem by separation of variables,[5] let $u(r,z) = R(r)Z(z)$ and

$$\frac{1}{rR}\frac{d}{dr}\left(r\frac{dR}{dr}\right) = -\frac{1}{Z}\frac{d^2Z}{dz^2} = -\frac{k^2}{a^2}. \tag{9.3.35}$$

Only a negative separation constant yields nontrivial solutions in the radial direction. In that case, we have that

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dR}{dr}\right) + \frac{k^2}{a^2}R = 0. \tag{9.3.36}$$

The solutions of Equation 9.3.36 are the Bessel functions $J_0(kr/a)$ and $Y_0(kr/\ a)$. Because $Y_0(kr/a)$ becomes infinite at $r = 0$, the only permissible solution is $J_0(kr/a)$. The condition that $u(a,z) = R(a)Z(z) = 0$ forces us to choose values of k such that $J_0(k) = 0$. Therefore, the solution in the radial direction is $J_0(k_n r/a)$, where k_n is the nth root of $J_0(k) = 0$.

In the z direction,

$$\frac{d^2Z_n}{dz^2} + \frac{k_n^2}{a^2}Z_n = 0. \tag{9.3.37}$$

The general solution to Equation 9.3.37 is

$$Z_n(z) = A_n \sinh\left(\frac{k_n z}{a}\right) + B_n \cosh\left(\frac{k_n z}{a}\right). \tag{9.3.38}$$

Because $u(r,0) = R(r)Z(0) = 0$ and $\cosh(0) = 1$, B_n must equal zero. Therefore, the general product solution is

$$u(r,z) = \sum_{n=1}^{\infty} A_n J_0\left(\frac{k_n r}{a}\right)\sinh\left(\frac{k_n z}{a}\right). \tag{9.3.39}$$

The condition that $u(r,L) = V$ determines the arbitrary constant A_n. Along $z = L$,

$$u(r,L) = V = \sum_{n=1}^{\infty} A_n J_0\left(\frac{k_n r}{a}\right)\sinh\left(\frac{k_n L}{a}\right), \tag{9.3.40}$$

where

$$\sinh\left(\frac{k_n L}{a}\right) A_n = \frac{2V}{a^2 J_1^2(k_n)} \int_0^L r\ J_0\left(\frac{k_n r}{a}\right)\,dr \tag{9.3.41}$$

from Equation 6.5.38 and Equation 6.5.45. Thus,

$$\sinh\left(\frac{k_n L}{a}\right) A_n = \frac{2V}{k_n^2 J_1^2(k_n)}\left(\frac{k_n r}{a}\right)\ J_1\left(\frac{k_n r}{a}\right)\Bigg|_0^a = \frac{2V}{k_n J_1(k_n)}. \tag{9.3.42}$$

The solution is then

$$u(r,z) = 2V\sum_{n=1}^{\infty} \frac{J_0(k_n r/a)}{k_n J_1(k_n)}\frac{\sinh(k_n z/a)}{\sinh(k_n L/a)}. \tag{9.3.43}$$

---

[5] Wang and Liu [Wang, M.-L., and B.-L. Liu, 1995: Solution of Laplace equation by the method of separation of variables. *J. Chinese Inst. Eng.*, **18**, 731–739] have written a review article on the solutions to Equation 9.3.33 based upon which order the boundary conditions are satisfied.

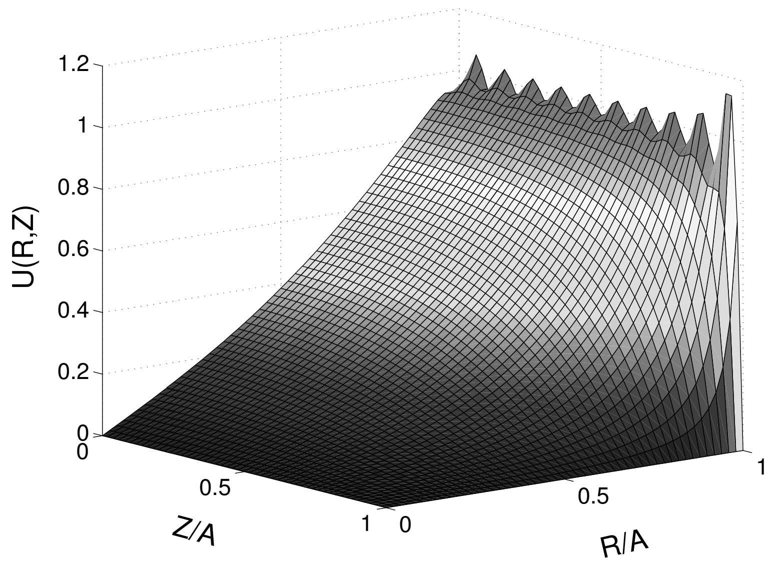


Figure 9.3.3: The steady-state potential (divided by V) within a cylinder of equal radius and height a when the top has the potential V while the lateral side and bottom are at potential 0.

Figure 9.3.3 illustrates Equation 9.3.43 for the case when $L = a$ where we included the first 20 terms of the series. It was created using the MATLAB script

```
clear
L_over_a = 1; M = 20; dr = 0.02; dz = 0.02;
% load in zeros of J_0
zero( 1) =  2.40482; zero( 2) =  5.52007; zero( 3) =  8.65372;
zero( 4) = 11.79153; zero( 5) = 14.93091; zero( 6) = 18.07106;
zero( 7) = 21.21164; zero( 8) = 24.35247; zero( 9) = 27.49347;
zero(10) = 30.63461; zero(11) = 33.77582; zero(12) = 36.91710;
zero(13) = 40.05843; zero(14) = 43.19979; zero(15) = 46.34119;
zero(16) = 49.48261; zero(17) = 52.62405; zero(18) = 55.76551;
zero(19) = 58.90698; zero(20) = 62.04847;
% compute Fourier coefficients
for m = 1:M
  a(m) = 2/(zero(m)*besselj(1,zero(m))*sinh(L_over_a*zero(m)));
end
% compute grid and initialize solution
R_over_a = [0:dr:1]; Z_over_a = [0:dz:1];
u = zeros(length(Z_over_a),length(R_over_a));
RR_over_a = repmat(R_over_a,[length(Z_over_a) 1]);
ZZ_over_a = repmat(Z_over_a',[1 length(R_over_a)]);
% compute solution from Equation 9.3.43
for m = 1:M
  u=u+a(m).*besselj(0,zero(m)*RR_over_a) .* sinh(zero(m)*ZZ_over_a);
end
surf(RR_over_a,ZZ_over_a,u)
xlabel('R/A','Fontsize',20); ylabel('Z/A','Fontsize',20)
zlabel('U(R,Z)','Fontsize',20)
```

Of particular interest are the ripples along the line $z = L$. Along that line, the solution must jump from V to 0 at $r = a$. For that reason our solution suffers from Gibbs phenomena along this boundary. As we move away from that region the electrostatic potential varies smoothly. □

• **Example 9.3.4**

Let us now consider a similar, but slightly different, version of Example 9.3.3, where the ends are held at zero potential while the lateral side has the value V. Once again, the governing equation is Equation 9.3.33 with the boundary conditions

$$u(r,0) = u(r,L) = 0, \quad \text{and} \quad u(a,z) = V. \tag{9.3.44}$$

Separation of variables yields

$$\frac{1}{rR}\frac{d}{dr}\left(r\frac{dR}{dr}\right) = -\frac{1}{Z}\frac{d^2Z}{dz^2} = \frac{k^2}{L^2} \tag{9.3.45}$$

with $Z(0) = Z(L) = 0$. We chose a positive separation constant because a negative constant would give hyperbolic functions in z that cannot satisfy the boundary conditions. A separation constant of zero would give a straight line for $Z(z)$. Applying the boundary conditions gives a trivial solution. Consequently, the only solution in the z direction that satisfies the boundary conditions is $Z_n(z) = \sin(n\pi z/L)$.

In the radial direction, the differential equation is

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dR_n}{dr}\right) - \frac{n^2\pi^2}{L^2}R_n = 0. \tag{9.3.46}$$

As we showed in Section 6.5, the general solution is

$$R_n(r) = A_n I_0\left(\frac{n\pi r}{L}\right) + B_n K_0\left(\frac{n\pi r}{L}\right), \tag{9.3.47}$$

where I_0 and K_0 are modified Bessel functions of the first and second kind, respectively, of order zero. Because $K_0(x)$ behaves as $-\ln(x)$ as $x \to 0$, we must discard it and our solution in the radial direction becomes $R_n(r) = A_n I_0(n\pi r/L)$. Hence, the product solution is

$$u_n(r,z) = A_n I_0\left(\frac{n\pi r}{L}\right)\sin\left(\frac{n\pi z}{L}\right), \tag{9.3.48}$$

and the general solution is a sum of these particular solutions, namely

$$u(r,z) = \sum_{n=1}^{\infty} A_n I_0\left(\frac{n\pi r}{L}\right)\sin\left(\frac{n\pi z}{L}\right). \tag{9.3.49}$$

Finally, we use the boundary conditions that $u(a,z) = V$ to compute A_n. This condition gives

$$u(a,z) = V = \sum_{n=1}^{\infty} A_n I_0\left(\frac{n\pi a}{L}\right)\sin\left(\frac{n\pi z}{L}\right), \tag{9.3.50}$$

so that

$$I_0\left(\frac{n\pi a}{L}\right) A_n = \frac{2}{L}\int_0^L V\sin\left(\frac{n\pi z}{L}\right)\,dz = \frac{2V[1-(-1)^n]}{n\pi}. \tag{9.3.51}$$

Therefore, the final answer is

$$u(r,z) = \frac{4V}{\pi}\sum_{m=1}^{\infty}\frac{I_0[(2m-1)\pi r/L]\sin[(2m-1)\pi z/L]}{(2m-1)I_0[(2m-1)\pi a/L]}. \quad (9.3.52)$$

Figure 9.3.4 illustrates the solution, Equation 9.3.52, for the case when $L = a$. It was created using the MATLAB script

```
clear
a_over_L = 1; M = 200; dr = 0.02; dz = 0.02;
% compute grid and initialize solution
R_over_L = [0:dr:1]; Z_over_L = [0:dz:1];
u = zeros(length(Z_over_L),length(R_over_L));
RR_over_L = repmat(R_over_L,[length(Z_over_L) 1]);
ZZ_over_L = repmat(Z_over_L',[1 length(R_over_L)]);

for m = 1:M
  temp = (2*m-1)*pi; prod1 = temp*a_over_L;
% compute modified bessel functions in Equation 9.3.52
  for j = 1:length(Z_over_L); for i = 1:length(R_over_L);
    prod2 = temp*RR_over_L(i,j);
    if prod2 - prod1 > -10
      if prod2 < 20
        ratio(i,j) = besseli(0,prod2) / besseli(0,prod1);
      else
% for large values of prod, use asymptotic expansion
%          for modified bessel function
        ratio(i,j) = sqrt(prod1/prod2) * exp(prod2-prod1); end;
    else
      ratio(i,j) = 0; end
  end; end;
% compute solution from Equation 9.3.52
  u = u + (4/temp) * ratio .* sin(temp*ZZ_over_L);
end
surf(RR_over_L,ZZ_over_L,u)
xlabel('R/L','Fontsize',20); ylabel('Z/L','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

Once again, there is a convergence of equipotentials at the corners along the right side. If we had plotted more contours, we would have observed Gibbs phenomena in the solution along the top and bottom of the cylinder. □

• **Example 9.3.5**

In the previous examples, the domain was always of finite extent. Assuming axial symmetry, let us now solve Laplace's equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < \infty, \quad 0 < z < \infty, \quad (9.3.53)$$

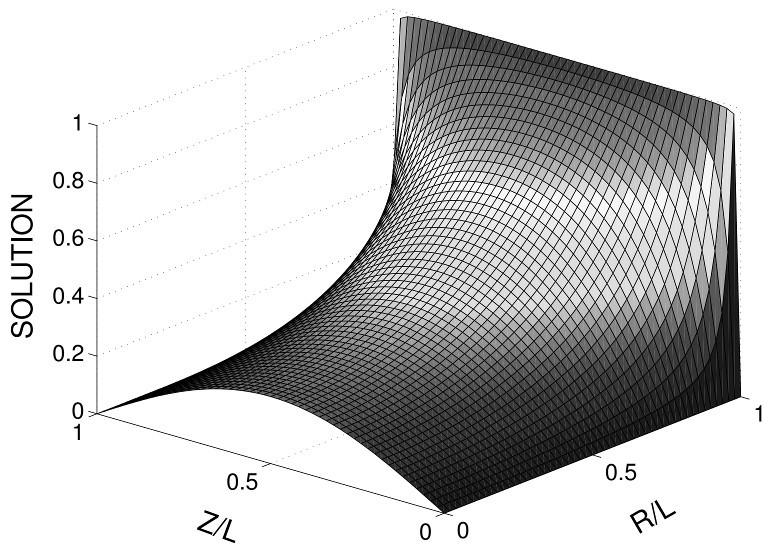


Figure 9.3.4: Potential (divided by V) within a conducting cylinder when the top and bottom have a potential 0 while the lateral side has a potential V.

in the half-plane $z > 0$ subject to the boundary conditions

$$\lim_{z\to\infty} |u(r,z)| < \infty, \qquad u(r,0) = \begin{cases} u_0, & r < a, \\ 0, & r > a, \end{cases} \tag{9.3.54}$$

$$\lim_{r\to 0} |u(r,z)| < \infty, \qquad \text{and} \qquad \lim_{r\to\infty} |u(r,z)| < \infty. \tag{9.3.55}$$

This problem gives the steady-state temperature distribution in the half-space $z > 0$ where the temperature on the bounding plane $z = 0$ equals u_0 within a circle of radius a and equals 0 outside of the circle.

As before, we begin by assuming the product solution $u(r,z) = R(r)Z(z)$ and separate the variables. Again, the separation constant may be positive, negative, or zero. Turning to the positive separation constant first, we have that

$$\frac{R''}{R} + \frac{1}{r}\frac{R'}{R} = -\frac{Z''}{Z} = m^2. \tag{9.3.56}$$

Focusing on the R equation,

$$\frac{R''}{R} + \frac{1}{r}\frac{R'}{R} - m^2 = 0, \quad \text{or} \quad r^2R'' + rR - m^2r^2R = 0. \tag{9.3.57}$$

The solution to Equation 9.3.57 is

$$R(r) = A_1 I_0(mr) + A_2 K_0(mr), \tag{9.3.58}$$

where $I_0(\cdot)$ and $K_0(\cdot)$ denote modified Bessel functions of order zero and the first and second kind, respectively. Because $u(r,z)$, and hence $R(r)$, must be bounded as $r \to 0$, $A_2 = 0$. Similarly, since $u(r,z)$ must also be bounded as $r \to \infty$, $A_1 = 0$ because $\lim_{r\to\infty} I_0(mr) \to \infty$. Thus, there is only a trivial solution for a positive separation constant.

We next try the case when the separation constant equals 0. This yields

$$\frac{R''}{R} + \frac{1}{r}\frac{R'}{R} = 0, \quad \text{or} \quad r^2R'' + rR = 0. \tag{9.3.59}$$

The solution here is

$$R(r) = A_1 + A_2 \ln(r). \tag{9.3.60}$$

Again, boundedness as $r \to 0$ requires that $A_2 = 0$. What about A_1? Clearly, for any arbitrary value of z, the amount of internal energy must be finite. This corresponds to

$$\int_0^\infty |u(r,z)|\, dr < \infty \quad \text{or} \quad \int_0^\infty |R(r)|\, dr < \infty \tag{9.3.61}$$

and $A_1 = 0$. The choice of the zero separation constant yields a trivial solution.

Finally, when the separation constant equals $-k^2$, the equations for $R(r)$ and $Z(z)$ are

$$r^2R'' + rR + k^2r^2R = 0, \qquad \text{and} \qquad Z'' - k^2Z = 0, \tag{9.3.62}$$

respectively. Solving for $R(r)$ first, we have that

$$R(r) = A_1J_0(kr) + A_2Y_0(kr), \tag{9.3.63}$$

where $J_0(\cdot)$ and $Y_0(\cdot)$ denote Bessel functions of order zero and the first and second kind, respectively. The requirement that $u(r,z)$, and hence $R(r)$, is bounded as $r \to 0$ forces us to take $A_2 = 0$, leaving $R(r) = A_1J_0(kr)$. From the equation for $Z(z)$, we conclude that

$$Z(z) = B_1e^{kz} + B_2e^{-kz}. \tag{9.3.64}$$

Since $u(r,z)$, and hence $Z(z)$, must be bounded as $z \to \infty$, it follows that $B_1 = 0$, leaving $Z(z) = B_2e^{-kz}$.

Presently our analysis follows closely those for a finite domain. However, we have satisfied all of the boundary conditions and yet there is still *no* restriction on k. Consequently, we conclude that k is completely arbitrary and any product solution

$$u_k(r,z) = A_1B_2\, J_0(kr)\, e^{-kz} \tag{9.3.65}$$

is a solution to our partial differential equation and satisfies the boundary conditions. From the principle of linear superposition, the most general solution equals the sum of *all* of the possible solutions, or

$$u(r,z) = \int_0^\infty A(k)\, k\, J_0(kr)\, e^{-kz}\, dk, \tag{9.3.66}$$

where we have written the arbitrary constant A_1B_2 as $A(k)k$. Our final task remains to compute $A(k)$.

Before we can find $A(k)$, we must derive an intermediate result. If we define our Fourier transform in an appropriate manner, we can write the two-dimensional Fourier transform pair as

$$f(x,y) = \frac{1}{2\pi}\int_{-\infty}^\infty \int_{-\infty}^\infty F(k,\ell)\, e^{i(kx+\ell y)}\, dk\, d\ell, \tag{9.3.67}$$

where

$$F(k,\ell) = \frac{1}{2\pi}\int_{-\infty}^\infty \int_{-\infty}^\infty f(x,y)\, e^{-i(kx+\ell y)}\, dx\, dy. \tag{9.3.68}$$

Consider now the special case where $f(x,y)$ is only a function of $r = \sqrt{x^2+y^2}$, so that $f(x,y) = g(r)$. Then, changing to polar coordinates through the substitution $x = r\cos(\theta)$, $y = r\sin(\theta)$, $k = \rho\cos(\varphi)$, and $\ell = \rho\sin(\varphi)$, we have that

$$kx + \ell y = r\rho[\cos(\theta)\cos(\varphi) + \sin(\theta)\sin(\varphi)] = r\rho\cos(\theta - \varphi), \tag{9.3.69}$$

and

$$dA = dx\, dy = r\, dr\, d\theta. \tag{9.3.70}$$

Therefore, the integral in Equation 9.3.68 becomes

$$F(k,\ell) = \frac{1}{2\pi}\int_0^\infty \int_0^{2\pi} g(r)\, e^{-ir\rho\cos(\theta-\varphi)} r\, dr\, d\theta \tag{9.3.71}$$

$$= \frac{1}{2\pi}\int_0^\infty r\, g(r) \left[\int_0^{2\pi} e^{-ir\rho\cos(\theta-\varphi)}\, d\theta\right] dr. \tag{9.3.72}$$

If we introduce $\lambda = \theta - \varphi$, the integral

$$\int_0^{2\pi} e^{-ir\rho\cos(\theta-\varphi)}\, d\theta = \int_{-\varphi}^{2\pi-\varphi} e^{-ir\rho\cos(\lambda)}\, d\lambda = \int_0^{2\pi} e^{-ir\rho\cos(\lambda)}\, d\lambda = 2\pi J_0(\rho r). \tag{9.3.73}$$

The third integral in Equation 9.3.73 is equivalent to the second integral because the integral of a periodic function over one full period is the same regardless of where the integration begins. The final result follows from the integral definition of the Bessel function.[6] Therefore,

$$F(k,\ell) = \int_0^\infty r\, g(r)\, J_0(\rho r)\, dr. \tag{9.3.74}$$

Finally, because Equation 9.3.74 is clearly a function of $\rho = \sqrt{k^2+\ell^2}$, $F(k,\ell) = G(\rho)$ and

$$G(\rho) = \int_0^\infty r\, g(r)\, J_0(\rho r)\, dr. \tag{9.3.75}$$

Conversely, if we begin with Equation 9.3.67, make the same substitution, and integrate over the $k\ell$ plane, we have that

$$f(x,y) = g(r) = \frac{1}{2\pi}\int_0^\infty \int_0^{2\pi} F(k,\ell)\, e^{ir\rho\cos(\theta-\varphi)} \rho\, d\rho\, d\varphi \tag{9.3.76}$$

$$= \frac{1}{2\pi}\int_0^\infty \rho\, G(\rho) \left[\int_0^{2\pi} e^{ir\rho\cos(\theta-\varphi)}\, d\varphi\right] d\rho \tag{9.3.77}$$

$$= \int_0^\infty \rho\, G(\rho)\, J_0(\rho r)\, d\rho. \tag{9.3.78}$$

Thus, we obtain the result that if $\int_0^\infty |F(r)|\, dr$ exists, then

$$g(r) = \int_0^\infty \rho\, G(\rho)\, J_0(\rho r)\, d\rho, \tag{9.3.79}$$

where

$$G(\rho) = \int_0^\infty r\, g(r)\, J_0(\rho r)\, dr. \tag{9.3.80}$$

Taken together, Equation 9.3.79 and Equation 9.3.80 constitute the *Hankel transform pair for Bessel function of order 0*. The function $G(\rho)$ is called the Hankel transform of $g(r)$.

---

[6] Watson, G. N., 1966: *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, Section 2.2, Equation 5.

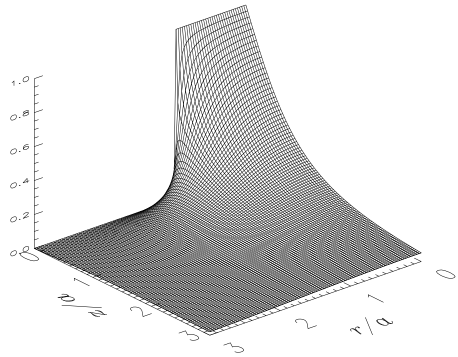


Figure 9.3.5: The axisymmetric potential $u(r,z)/u_0$ in the half-space $z>0$ when $u(r,0)=u_0$ if $r<a$ and $u(r,0)=0$ if $r>a$.

Why did we introduce Hankel transforms? First, setting $z=0$ in Equation 9.3.66, we find that

$$u(r,0) = \int_0^\infty A(k)\,k\,J_0(kr)\,dk. \qquad \textbf{(9.3.81)}$$

If we now compare Equation 9.3.79 with Equation 9.3.81, we recognize that $A(k)$ is the Hankel transform of $u(r,0)$. Therefore,

$$A(k) = \int_0^\infty r\,u(r,0)\,J_0(kr)\,dr = u_0 \int_0^a r\,J_0(kr)\,dr = \frac{u_0}{k}\,r\,J_1(kr)\Big|_0^a = \frac{au_0}{k}J_1(ka). \qquad \textbf{(9.3.82)}$$

Thus, the complete solution is

$$u(r,z) = au_0 \int_0^\infty J_1(ka)\,J_0(kr)\,e^{-kz}\,dk. \qquad \textbf{(9.3.83)}$$

Equation 9.3.83 is illustrated in Figure 9.3.5. □

• Example 9.3.6: Mixed boundary-value problem

In all of our previous examples, the boundary condition along any specific boundary remained the same. In this example, we relax this condition and consider a *mixed boundary-value problem*.

Consider[7] the axisymmetric Laplace equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < \infty, \quad 0 < z < 1, \qquad \textbf{(9.3.84)}$$

subject to the boundary conditions

$$\lim_{r\to 0}|u(r,z)| < \infty, \qquad \lim_{r\to\infty}|u(r,z)| < \infty, \qquad u(r,0) = 0, \qquad \textbf{(9.3.85)}$$

and

$$\begin{cases} u(r,1) = 1, & 0 < r \le a, \\ u(r,1) + \dfrac{u_z(r,1)}{\sigma} = 1, & a < r < \infty. \end{cases} \qquad \textbf{(9.3.86)}$$

[7] See Nir, A., and R. Pfeffer, 1979: Transport of macromolecules across arterial wall in the presence of local endothial injury. *J. Theor. Biol.*, **81**, 685–711.

The interesting aspect of this example is the mixture of boundary conditions along the boundary $z = 1$. For $r \leq a$, we have a Dirichlet boundary condition, which becomes a Robin boundary condition when $r > a$.

Our analysis begins as it did in the previous examples with separation of variables and a superposition of solutions. In the present case the solution is

$$u(r, z) = \frac{\sigma z}{1 + \sigma} + \frac{a}{1 + \sigma} \int_0^\infty A(k, a) \sinh(kz) J_0(kr)\, dk. \tag{9.3.87}$$

The first term on the right side of Equation 9.3.87 arises from a separation constant that equals zero while the second term is the contribution from a negative separation constant. Note that this equation satisfies all of the boundary conditions given in Equation 9.3.85. Substitution of Equation 9.3.87 into Equation 9.3.86 leads to the *dual integral equations*:

$$a \int_0^\infty A(k, a) \sinh(k) J_0(kr)\, dk = 1, \tag{9.3.88}$$

if $0 < r \leq a$, and

$$\int_0^\infty A(k, a) \left[\sinh(k) + \frac{k \cosh(k)}{\sigma} \right] J_0(kr)\, dk = 0, \tag{9.3.89}$$

if $a < r < \infty$.

What sets this problem apart from the routine separation of variables is the solution of dual integral equations;[8] in general, they are very difficult to solve. The process usually begins with finding a solution that satisfies Equation 9.3.89 via the orthogonality condition involving Bessel functions. This is the technique employed by Tranter,[9] who proved that the dual integral equations:

$$\int_0^\infty G(\lambda) f(\lambda) J_0(\lambda a)\, d\lambda = g(a), \tag{9.3.90}$$

and

$$\int_0^\infty f(\lambda) J_0(\lambda a)\, d\lambda = 0 \tag{9.3.91}$$

have the solution

$$f(\lambda) = \lambda^{1-\kappa} \sum_{n=0}^\infty A_n J_{2m+\kappa}(\lambda), \tag{9.3.92}$$

if $G(\lambda)$ and $g(a)$ are known. The value of κ is chosen so that the difference $G(\lambda) - \lambda^{2\kappa-2}$ is fairly small. In the present case, $f(\lambda) = \sinh(\lambda) A(\lambda, a)$, $g(a) = 1$, and $G(\lambda) = 1 + \lambda \coth(\lambda)/\sigma$.

What is the value of κ here? Clearly we would like our solution to be valid for a wide range of σ. Because $G(\lambda) \to 1$ as $\sigma \to \infty$, a reasonable choice is $\kappa = 1$. Therefore, we take

$$\sinh(k) A(k, a) = \sum_{n=1}^\infty \frac{A_n}{1 + k \coth(k)/\sigma} J_{2n-1}(ka). \tag{9.3.93}$$

---

[8] The standard reference is Sneddon, I. N., 1966: *Mixed Boundary Value Problems in Potential Theory*. Wiley, 283 pp.

[9] Tranter, C. J., 1950: On some dual integral equations occurring in potential problems with axial symmetry. *Quart. J. Mech. Appl. Math.*, **3**, 411–419.

Our final task remains to find A_n.

We begin by writing

$$\frac{A_n}{1 + k\coth(k)/\sigma} J_{2n-1}(ka) = \sum_{m=1}^{\infty} B_{mn}\, J_{2m-1}(ka), \tag{9.3.94}$$

where B_{mn} depends only on a and σ. Multiplying Equation 9.3.94 by the factor $J_{2p-1}(ka)$ dk/k and integrating,

$$\int_0^\infty \frac{A_n}{1 + k\coth(k)/\sigma}\, J_{2n-1}(ka)\, J_{2p-1}(ka)\, \frac{dk}{k} = \int_0^\infty \sum_{m=1}^{\infty} B_{nm}\, J_{2m-1}(ka)\, J_{2p-1}(ka) \frac{dk}{k}. \tag{9.3.95}$$

Because[10]

$$\int_0^\infty J_{2n-1}(ka)\, J_{2p-1}(ka)\, \frac{dk}{k} = \frac{\delta_{mp}}{2(2m-1)}, \tag{9.3.96}$$

where δ_{mp} is the Kronecker delta:

$$\delta_{mp} = \begin{cases} 1, & m = p, \\ 0, & m \neq p, \end{cases} \tag{9.3.97}$$

Equation 9.3.95 reduces to

$$A_n \int_0^\infty \frac{J_{2n-1}(ka) J_{2p-1}(ka)}{1 + k\coth(k)/\sigma}\, \frac{dk}{k} = \frac{B_{mn}}{2(2m-1)}. \tag{9.3.98}$$

If we define

$$\int_0^\infty \frac{J_{2n-1}(ka)\, J_{2m-1}(ka)}{1 + k\coth(k)/\sigma}\, \frac{dk}{k} = S_{mn}, \tag{9.3.99}$$

then we can rewrite Equation 9.3.98 as

$$A_n S_{mn} = \frac{B_{mn}}{2(2m-1)}. \tag{9.3.100}$$

Because[11]

$$a \int_0^\infty J_0(kr)\, J_{2m-1}(ka)\, dk = P_{m-1}\left(1 - \frac{2r^2}{a^2}\right), \tag{9.3.101}$$

if $r < a$, where $P_m(\ )$ is the Legendre polynomial of order m, Equation 9.3.88 can be rewritten

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} B_{mn} P_{m-1}\left(1 - \frac{2r^2}{a^2}\right) = 1. \tag{9.3.102}$$

Equation 9.3.102 follows from the substitution of Equation 9.3.93 into Equation 9.3.88 and then using Equation 9.3.101. Multiplying Equation 9.3.102 by $P_{m-1}(\xi)\, d\xi$, integrating

---

[10] Gradshteyn, I. S., and I. M. Ryzhik, 1965: *Table of Integrals, Series, and Products*. Academic Press, Section 6.538, Formula 2.

[11] Ibid., Section 6.512, Formula 4.

Table 9.3.1: The Convergence of the Coefficients A_n Given by Equation 9.3.104 Where S_{mn} Has Nonzero Values for $1 \le m, n \le N$

N	A_1	$-A_2$	A_3	$-A_4$	A_5	$-A_6$	A_7	$-A_8$
1	2.9980							
2	3.1573	1.7181						
3	3.2084	2.0329	1.5978					
4	3.2300	2.1562	1.9813	1.4517				
5	3.2411	2.2174	2.1548	1.8631	1.3347			
6	3.2475	2.2521	2.2495	2.0670	1.7549	1.2399		
7	3.2515	2.2738	2.3073	2.1862	1.9770	1.6597	1.1620	
8	3.2542	2.2882	2.3452	2.2626	2.1133	1.8925	1.5772	1.0972

between -1 and 1, and using the orthogonality properties of the Legendre polynomial, we have that

$$\sum_{n=1}^{\infty} B_{mn} \int_{-1}^{1} [P_{m-1}(\xi)]^2 \, d\xi = \int_{-1}^{1} P_{m-1}(\xi) \, d\xi = \int_{-1}^{1} P_0(\xi) P_{m-1}(\xi) \, d\xi, \qquad \textbf{(9.3.103)}$$

which shows that only $m = 1$ yields a nontrivial sum. Thus,

$$\sum_{n=1}^{\infty} B_{mn} = 2(2m-1) \sum_{n=1}^{\infty} A_n S_{mn} = 0, \quad m \ge 2, \qquad \textbf{(9.3.104)}$$

and

$$\sum_{n=1}^{\infty} B_{1n} = 2 \sum_{n=1}^{\infty} A_n S_{1n} = 1, \qquad \text{or} \qquad \sum_{n=1}^{\infty} S_{mn} A_n = \tfrac{1}{2} \delta_{m1}. \qquad \textbf{(9.3.105)}$$

Thus, we have reduced the problem to the solution of an infinite number of linear equations which yield A_n — a common occurrence in the solution of dual integral equations. Selecting some maximum value for n and m, say N, each term in the matrix S_{mn}, $1 \le m, n \le N$, is evaluated numerically for a given value of a and σ. By inverting Equation 9.3.104, we obtain the coefficients A_n for $n = 1, \ldots, N$. Because we solved a truncated version of Equation 9.3.105, they will only be approximate. To find more accurate values, we can increase N by 1 and again invert Equation 9.3.104. In addition to the new A_{N+1}, the previous coefficients will become more accurate. We can repeat this process of increasing N until the coefficients converge to their correct value. This is illustrated in Table 9.3.1 when $\sigma = a = 1$.

Once we have computed the coefficients A_n necessary for the desired accuracy, we use Equation 9.3.93 to find $A(k, a)$ and then obtain $u(r, z)$ from Equation 9.3.87 via numerical integration. Figure 9.3.6 illustrates the solution when $\sigma = 1$ and $a = 2$.

Mixed boundary-value problems over a finite domain can be solved in a similar manner. Consider the partial differential equation[12]

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < a, \quad 0 < z < 1, \qquad \textbf{(9.3.106)}$$

[12] See Vrentas, J. S., D. C. Venerus, and C. M. Vrentes, 1991: An exact analysis of reservoir effects for rotational viscometers. *Chem. Eng. Sci.*, **46**, 33–37.

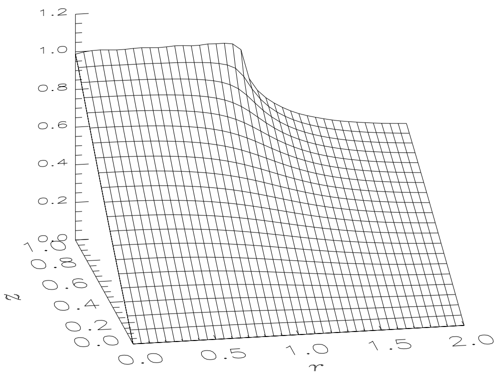


Figure 9.3.6: The solution of the axisymmetric Laplace's equation, Equation 9.3.84, with $u(r,0)=0$ and the mixed boundary condition, Equation 9.3.86. Here we have chosen $\sigma=1$ and $a=2$.

subject to the boundary conditions

$$\lim_{r\to 0}|u(r,z)|<\infty, \quad u(a,z)=0, \quad 0\le z\le 1, \tag{9.3.107}$$

and

$$u(r,0)=0, \quad 0\le r\le a, \quad \begin{cases} u(r,1)=r, & 0\le r<1, \\ u_z(r,1)=0, & 1<r\le a. \end{cases} \tag{9.3.108}$$

We begin by solving Equation 9.3.106 via separation of variables. This yields

$$u(r,z)=\sum_{n=1}^{\infty} A_n \sinh(k_n z)\, J_1(k_n r), \tag{9.3.109}$$

where k_n is the nth root of $J_1(ka)=0$. Note that Equation 9.3.109 satisfies all of the boundary conditions except those along $z=1$. Substituting Equation 9.3.109 into Equation 9.3.108, we find that

$$\sum_{n=1}^{\infty} A_n \sinh(k_n)\, J_1(k_n r)=r, \quad 0\le r<1, \tag{9.3.110}$$

and

$$\sum_{n=1}^{\infty} k_n A_n \cosh(k_n)\, J_1(k_n r)=0, \quad 1<r\le a. \tag{9.3.111}$$

Equations 9.3.110 and 9.3.111 show that in place of dual integral equations, we now have *dual Fourier-Bessel series*. Cooke and Tranter[13] have shown that the dual Fourier-Bessel series

$$\sum_{n=1}^{\infty} a_n J_\nu(k_n r)=0, \quad 1<r<a, \quad -1<\nu, \tag{9.3.112}$$

Other examples include:

Sherwood, J. D., and H. A. Stone, 1997: Added mass of a disc accelerating within a pipe. *Phys. Fluids*, **9**, 3141–3148.

Galceran, J., J. Cecília, E. Companys, J. Salvador, and J. Puy, 2000: Analytical expressions for feedback currents at the scanning electrochemical microscope. *J. Phys. Chem. B*, **104**, 7993–8000.

[13] Cooke, J. C., and C. J. Tranter, 1959: Dual Fourier-Bessel series. *Quart. J. Mech. Appl. Math.*, **12**, 379–386.

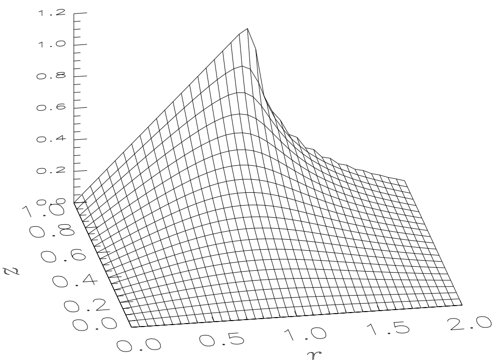


Figure 9.3.7: The solution of Equation 9.3.106, which satisfies the boundary condition, Equation 9.3.107, and the mixed boundary condition, Equation 9.3.108. Here we have chosen $a = 2$.

where $J_\nu(k_n a) = 0$, will be automatically satisfied if

$$k_n^{1+p/2} J_{\nu+1}^2(k_n a) a_n = \sum_{m=0}^{\infty} b_m J_{\nu+2m+1+p/2}(k_m), \qquad \textbf{(9.3.113)}$$

where $|p| \le 1$. Because $a_n = k_n A_n \cosh(k_n)$ and $\nu = 1$ here, A_n is given by

$$k_n^2 \cosh(k_n) J_2^2(k_n a) A_n = \sum_{m=1}^{\infty} B_m J_{2m}(k_n), \qquad \textbf{(9.3.114)}$$

if we take $p = 0$.

Substitution of Equation 9.3.114 into Equation 9.3.110 gives

$$\sum_{m=1}^{\infty} B_m \sum_{n=1}^{\infty} \frac{\sinh(k_n) J_{2m}(k_n) J_1(k_n r)}{k_n^2 \cosh(k_n) J_2^2(k_n a)} = r. \qquad \textbf{(9.3.115)}$$

Multiplying both sides of this equation by $r J_1(k_p r)\, dr$, $p = 1, 2, 3, \ldots$, and integrating from 0 to 1, we find that

$$\sum_{m=1}^{\infty} B_m \sum_{n=1}^{\infty} \frac{\sinh(k_n) J_{2m}(k_n) Q_{pn}}{k_n^2 \cosh(k_n) J_2^2(k_n a)} = \int_0^1 r^2 J_p(k_p r)\, dr, \qquad \textbf{(9.3.116)}$$

where

$$Q_{pn} = \int_0^1 J_1(k_p r) J_1(k_n r)\, r\, dr = \begin{cases} \frac{k_p J_1(k_n) J_0(k_p) - k_n J_1(k_p) J_0(k_n)}{k_n^2 - k_p^2}, & n \ne p, \\ \frac{J_1^2(k_p) - J_0(k_p) J_2(k_p)}{2}, & n = p. \end{cases} \qquad \textbf{(9.3.117)}$$

Carrying out the integration, Equation 9.3.116 yields the infinite set of equations

$$\sum_{m=1}^{\infty} M_{pm} B_m = \frac{J_2(k_p)}{k_p}, \qquad \textbf{(9.3.118)}$$

where

$$M_{pm} = \sum_{n=1}^{\infty} \frac{\sinh(k_n) J_{2m}(k_n) Q_{pn}}{k_n^2 \cosh(k_n) J_2^2(k_n a)}. \qquad \textbf{(9.3.119)}$$

Once again, we compute B_m by truncating Equation 9.3.118 to M terms and inverting the systems of equations. Increasing the value of M yields more accurate results. Once we have B_m, we use Equation 9.3.114 to find A_n. Finally, $u(r, z)$ follows from Equation 9.3.109. Figure 9.3.7 illustrates $u(r, z)$ when $a = 2$. □

• Example 9.3.7

Let us find the potential at any point P within a conducting sphere of radius a. At the surface, the potential is held at V_0 in the hemisphere $0 < \theta < \pi/2$, and $-V_0$ for $\pi/2 < \theta < \pi$.

Laplace's equation in spherical coordinates is

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}\left[\sin(\theta)\frac{\partial u}{\partial \theta}\right] = 0, \quad 0 \le r < a, \quad 0 \le \theta \le \pi. \qquad \textbf{(9.3.120)}$$

To solve Equation 9.3.120 we set $u(r,\theta) = R(r)\Theta(\theta)$ by separation of variables. Substituting into this equation, we have that

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) = -\frac{1}{\sin(\theta)\Theta}\frac{d}{d\theta}\left[\sin(\theta)\frac{d\Theta}{d\theta}\right] = k^2, \qquad \textbf{(9.3.121)}$$

or

$$r^2R'' + 2rR' - k^2R = 0, \qquad \textbf{(9.3.122)}$$

and

$$\frac{1}{\sin(\theta)}\frac{d}{d\theta}\left[\sin(\theta)\frac{d\Theta}{d\theta}\right] + k^2\Theta = 0. \qquad \textbf{(9.3.123)}$$

A common substitution replaces θ with $\mu = \cos(\theta)$. Then, as θ varies from 0 to π, μ varies from 1 to -1. With this substitution, Equation 9.3.123 becomes

$$\frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] + k^2\Theta = 0. \qquad \textbf{(9.3.124)}$$

This is Legendre's equation, which we examined in Section 6.4. Consequently, because the solution must remain finite at the poles, $k^2 = n(n+1)$, and

$$\Theta_n(\theta) = P_n(\mu) = P_n[\cos(\theta)], \qquad \textbf{(9.3.125)}$$

where $n = 0, 1, 2, 3, \ldots$.

Turning to Equation 9.3.122, this equation is the equidimensional or Euler-Cauchy linear differential equation. One method of solving this equation consists of introducing a new independent variable s so that $r = e^s$, or $s = \ln(r)$. Because

$$\frac{d}{dr} = \frac{ds}{dr}\frac{d}{ds} = e^{-s}\frac{d}{ds}, \qquad \textbf{(9.3.126)}$$

it follows that

$$\frac{d^2}{dr^2} = \frac{d}{dr}\left(e^{-s}\frac{d}{ds}\right) = e^{-s}\frac{d}{ds}\left(e^{-s}\frac{d}{ds}\right) = e^{-2s}\left(\frac{d^2}{ds^2} - \frac{d}{ds}\right). \qquad \textbf{(9.3.127)}$$

Substituting into Equation 9.3.122,

$$\frac{d^2R_n}{ds^2} + \frac{dR_n}{ds} - n(n+1)R_n = 0. \qquad \textbf{(9.3.128)}$$

Equation 9.3.128 is a second-order, constant coefficient ordinary differential equation, which has the solution

$$R_n(s) = C_ne^{ns} + D_ne^{-(n+1)s} = C_n\exp[n\ln(r)] + D_n\exp[-(n+1)\ln(r)] \qquad \textbf{(9.3.129)}$$

$$= C_n\exp[\ln(r^n)] + D_n\exp[\ln(r^{-1-n})] = C_nr^n + D_nr^{-1-n}. \qquad \textbf{(9.3.130)}$$

A more convenient form of the solution is

$$R_n(r) = A_n \left(\frac{r}{a}\right)^n + B_n \left(\frac{r}{a}\right)^{-1-n}, \tag{9.3.131}$$

where $A_n = a^n C_n$ and $B_n = D_n/a^{n+1}$. We introduced the constant a, the radius of the sphere, to simplify future calculations.

Using the results from Equation 9.3.125 and Equation 9.3.130, the solution to Laplace's equation in axisymmetric problems is

$$u(r,\theta) = \sum_{n=0}^{\infty} \left[A_n \left(\frac{r}{a}\right)^n + B_n \left(\frac{r}{a}\right)^{-1-n}\right] P_n[\cos(\theta)]. \tag{9.3.132}$$

In our particular problem we must take $B_n = 0$ because the solution becomes infinite at $r = 0$ otherwise. If the problem had involved the domain $a < r < \infty$, then $A_n = 0$ because the potential must remain finite as $r \to \infty$.

Finally, we must evaluate A_n. Finding the potential at the surface,

$$u(a,\mu) = \sum_{n=0}^{\infty} A_n P_n(\mu) = \begin{cases} V_0, & 0 < \mu \le 1, \\ -V_0, & -1 \le \mu < 0. \end{cases} \tag{9.3.133}$$

Upon examining Equation 9.3.133, it is merely an expansion in Legendre polynomials of the function

$$f(\mu) = \begin{cases} V_0, & 0 < \mu \le 1, \\ -V_0, & -1 \le \mu < 0. \end{cases} \tag{9.3.134}$$

Consequently, from Equation 9.3.133,

$$A_n = \frac{2n+1}{2} \int_{-1}^{1} f(\mu)\, P_n(\mu)\, d\mu. \tag{9.3.135}$$

Because $f(\mu)$ is an odd function, $A_n = 0$ if n is even. When n is odd, however,

$$A_n = (2n+1) \int_0^1 V_0\, P_n(\mu)\, d\mu. \tag{9.3.136}$$

We can further simplify Equation 9.3.136 by using the relationship that

$$\int_x^1 P_n(t)\, dt = \frac{1}{2n+1} \left[P_{n-1}(x) - P_{n+1}(x)\right], \tag{9.3.137}$$

where $n \ge 1$. In our problem, then,

$$A_n = \begin{cases} V_0[P_{n-1}(0) - P_{n+1}(0)], & n \text{ odd}, \\ 0, & n \text{ even}. \end{cases} \tag{9.3.138}$$

The first few terms are $A_1 = 3V_0/2$, $A_3 = -7V_0/8$, and $A_5 = 11V_0/16$.

Figure 9.3.8 illustrates our solution. It was created using the MATLAB script

```
clear
N = 51; dr = 0.05; dtheta = pi / 15;
% compute grid and set solution equal to zero
r = [0:dr:1]; theta = [0:dtheta:2*pi];
```

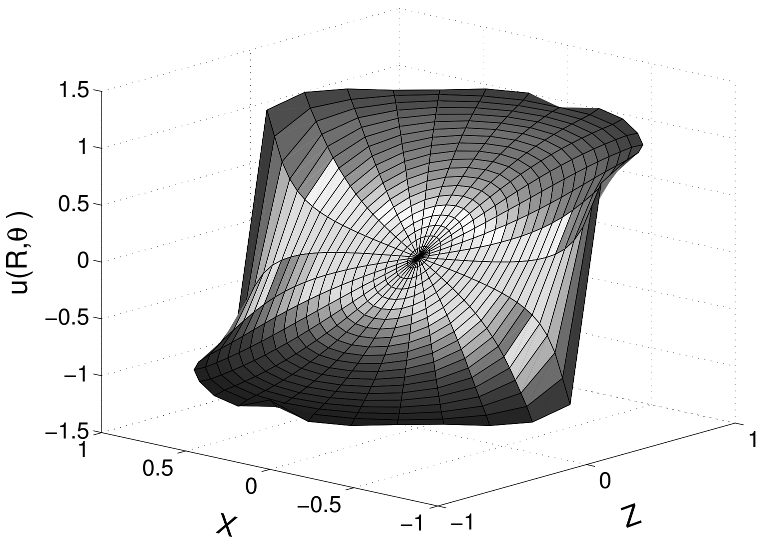


Figure 9.3.8: Electrostatic potential within a conducting sphere when the upper hemispheric surface has the potential 1 and the lower surface has the potential −1.

```
mu = cos(theta); Z = r' * mu;
for L = 1:2
  if L == 1 X = r' * sin(theta);
  else X = -r' * sin(theta); end
  u = zeros(size(X));
% compute solution from Equation 9.3.132
  rfactor = r;
  for n = 1:2:N
    A = legendre(n-1,0); B = legendre(n+1,0); coeff = A(1)-B(1);
    C = legendre(n,mu); Theta = C(1,:);
    u = u + coeff * rfactor' * Theta;
    rfactor = rfactor .* r .* r;
  end
  surf(Z,X,u); hold on; end
xlabel('Z','Fontsize',20); ylabel('X','Fontsize',20)
zlabel('u(R,\theta )','Fontsize',20);
```

Here we have the convergence of the equipotentials along the equator and at the surface. The slow rate at which the coefficients are approaching zero suggests that the solution suffers from Gibbs phenomena along the surface. □

• Example 9.3.8

We now find the steady-state temperature field within a metallic sphere of radius a, which we place in direct sunlight and allow to radiatively cool. This classic problem, first solved by Rayleigh,[14] requires the use of spherical coordinates with its origin at the center

[14] Rayleigh, J. W., 1870: On the values of the integral $\int_0^1 Q_n Q_{n'}\,d\mu$, Q_n, $Q_{n'}$ being Laplace's coefficients of the orders n, n', with application to the theory of radiation. *Philos. Trans. R. Soc. London, Ser. A*, **160**, 579–590.

of the sphere and its z-axis pointing toward the sun. With this choice for the coordinate system, the incident sunlight is

$$D(\theta) = \begin{cases} D(0)\cos(\theta), & 0 \le \theta \le \pi/2, \\ 0, & \pi/2 \le \theta \le \pi. \end{cases} \tag{9.3.139}$$

If heat dissipation takes place at the surface $r = a$ according to Newton's law of cooling and the temperature of the surrounding medium is zero, the solar heat absorbed by the surface dA must balance the Newtonian cooling at the surface plus the energy absorbed into the sphere's interior. This physical relationship is

$$(1-\rho)D(\theta)\,dA = \epsilon u(a,\theta)\,dA + \kappa \frac{\partial u(a,\theta)}{\partial r}\,dA, \tag{9.3.140}$$

where ρ is the reflectance of the surface (the albedo), ϵ is the surface conductance or coefficient of surface heat transfer, and κ is the thermal conductivity. Simplifying Equation 9.3.140, we have that

$$\frac{\partial u(a,\theta)}{\partial r} = \frac{1-\rho}{\kappa}D(\theta) - \frac{\epsilon}{\kappa}u(a,\theta) \tag{9.3.141}$$

for $r = a$.

If the sphere has reached thermal equilibrium, Laplace's equation describes the temperature field within the sphere. In the previous example, we showed that the solution to Laplace's equation in axisymmetric problems is

$$u(r,\theta) = \sum_{n=0}^{\infty} \left[A_n \left(\frac{r}{a}\right)^n + B_n \left(\frac{r}{a}\right)^{-1-n} \right] P_n[\cos(\theta)]. \tag{9.3.142}$$

In this problem, $B_n = 0$ because the solution would become infinite at $r = 0$ otherwise. Therefore,

$$u(r,\theta) = \sum_{n=0}^{\infty} A_n \left(\frac{r}{a}\right)^n P_n[\cos(\theta)]. \tag{9.3.143}$$

Differentiation gives

$$\frac{\partial u}{\partial r} = \sum_{n=0}^{\infty} A_n \frac{nr^{n-1}}{a^n} P_n[\cos(\theta)]. \tag{9.3.144}$$

Substituting into the boundary condition leads to

$$\sum_{n=0}^{\infty} A_n \left(\frac{n}{a} + \frac{\epsilon}{\kappa}\right) P_n[\cos(\theta)] = \left(\frac{1-\rho}{\kappa}\right) D(\theta), \tag{9.3.145}$$

or

$$D(\mu) = \sum_{n=0}^{\infty} \left[\frac{n\kappa + \epsilon a}{a(1-\rho)}\right] A_n P_n(\mu) = \sum_{n=0}^{\infty} C_n P_n(\mu), \tag{9.3.146}$$

where

$$C_n = \left[\frac{n\kappa + \epsilon a}{a(1-\rho)}\right] A_n, \quad \text{and} \quad \mu = \cos(\theta). \tag{9.3.147}$$

We determine the coefficients by

$$C_n = \frac{2n+1}{2} \int_{-1}^{1} D(\mu) P_n(\mu)\,d\mu = \frac{2n+1}{2} D(0) \int_0^1 \mu P_n(\mu)\,d\mu. \tag{9.3.148}$$

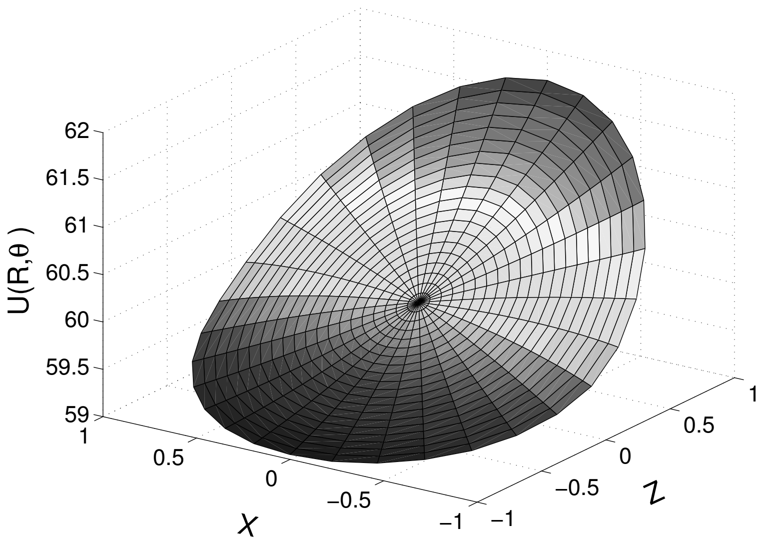


Figure 9.3.9: The difference (in °C) between the temperature field within a blackened iron surface of radius 0.1 m and the surrounding medium when we heat the surface by sunlight and allow it to radiatively cool.

Evaluation of the first few coefficients gives

$$A_0 = \frac{(1-\rho)D(0)}{4\epsilon}, \quad A_1 = \frac{a(1-\rho)D(0)}{2(\kappa+\epsilon a)}, \quad A_2 = \frac{5a(1-\rho)D(0)}{16(2\kappa+\epsilon a)}, \quad A_3 = 0, \tag{9.3.149}$$

$$A_4 = -\frac{3a(1-\rho)D(0)}{32(4\kappa+\epsilon a)}, \quad A_5 = 0, \quad A_6 = \frac{13a(1-\rho)D(0)}{256(6\kappa+\epsilon a)}, \quad A_7 = 0, \tag{9.3.150}$$

$$A_8 = -\frac{17a(1-\rho)D(0)}{512(8\kappa+\epsilon a)}, \quad A_9 = 0, \quad \text{and} \quad A_{10} = \frac{49a(1-\rho)D(0)}{2048(10\kappa+\epsilon a)}. \tag{9.3.151}$$

Figure 9.3.9 illustrates the temperature field within the sphere with $D(0) = 1200$ W/m^2, $\kappa = 45$ W/m K, $\epsilon = 5$ W/m^2 K, $\rho = 0$, and $a = 0.1$ m. This corresponds to a cast iron sphere with blackened surface in sunlight. This figure was created by the MATLAB script

```
clear
dr = 0.05; dtheta = pi / 15;
D_0 = 1200; kappa = 45; epsilon = 5; rho = 0; a = 0.1;
% compute grid and set solution equal to zero
r = [0:dr:1]; theta = [0:dtheta:pi];
mu = cos(theta); Z = r' * mu;
aaaa = (1-rho) * D_0 / ( 4 * epsilon);
aa(1) = a * (1-rho) * D_0 / ( 2 * ( kappa+epsilon*a));
aa(2) = 5 * a * (1-rho) * D_0 / ( 16 * (2*kappa+epsilon*a));
aa(3) = 0;
aa(4) = - 3 * a * (1-rho) * D_0 / ( 32 * (4*kappa+epsilon*a));
aa(5) = 0;
aa(6) = 13 * a * (1-rho) * D_0 / ( 256 * (6*kappa+epsilon*a));
aa(7) = 0;
aa(8) = -17 * a * (1-rho) * D_0 / ( 512 * (8*kappa+epsilon*a));
```

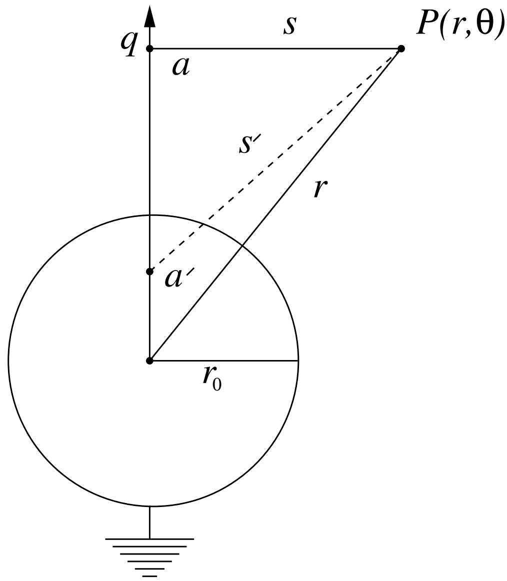


Figure 9.3.10: Point charge $+q$ in the presence of a grounded conducting sphere.

```
aa(9) = 0;
aa(10) = 49 * a * (1-rho) * D_0 / (2048 * (10*kappa+epsilon*a));
for L = 1:2
  if L == 1 X = r' * sin(theta);
  else X = -r' * sin(theta); end
  u = aaaa * ones(size(X));
  rfactor = r;
  for n = 1:10
    A = legendre(n,mu); Theta = A(1,:);
    u = u + aa(n) * rfactor' * Theta;
    rfactor = rfactor .* r;
  end
  surf(Z,X,u); hold on; end
xlabel('Z','Fontsize',20); ylabel('X','Fontsize',20);
zlabel('U(R,\theta )','Fontsize',20);
```

The temperature is quite warm with the highest temperature located at the position where the solar radiation is largest; the coolest temperatures are located in the shadow region. □

• **Example 9.3.9**

In this example we find the potential at any point P exterior to a conducting, grounded sphere centered at $z = 0$ after we place a point charge $+q$ at $z = a$ on the z-axis. See Figure 9.3.10. From the principle of linear superposition, the total potential $u(r, \theta)$ equals the sum of the potential from the point charge and the potential $v(r, \theta)$ due to the induced charge on the sphere

$$u(r, \theta) = \frac{q}{s} + v(r, \theta). \tag{9.3.152}$$

In common with the first term q/s, $v(r,\theta)$ must be a solution of Laplace's equation. In Example 9.3.7 we showed that the general solution to Laplace's equation in axisymmetric problems is

$$v(r,\theta) = \sum_{n=0}^{\infty}\left[A_n\left(\frac{r}{r_0}\right)^n + B_n\left(\frac{r}{r_0}\right)^{-1-n}\right]P_n[\cos(\theta)]. \tag{9.3.153}$$

Because the solutions must be valid *anywhere* outside of the sphere, $A_n = 0$; otherwise, the solution would not remain finite as $r \to \infty$. Hence,

$$v(r,\theta) = \sum_{n=0}^{\infty} B_n\left(\frac{r}{r_0}\right)^{-1-n} P_n[\cos(\theta)]. \tag{9.3.154}$$

We determine the coefficient B_n by the condition that $u(r_0,\theta) = 0$, or

$$\left.\frac{q}{s}\right|_{\text{on sphere}} + \sum_{n=0}^{\infty} B_n P_n[\cos(\theta)] = 0. \tag{9.3.155}$$

We need to expand the first term on the left side of Equation 9.3.155 in terms of Legendre polynomials. From the law of cosines,

$$s = \sqrt{r^2 + a^2 - 2ar\cos(\theta)}. \tag{9.3.156}$$

Consequently, if $a > r$, then

$$\frac{1}{s} = \frac{1}{a}\left[1 - 2\cos(\theta)\frac{r}{a} + \left(\frac{r}{a}\right)^2\right]^{-1/2}. \tag{9.3.157}$$

In Section 6.4, we showed that

$$(1 - 2xz + z^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(x) z^n. \tag{9.3.158}$$

Therefore,

$$\frac{1}{s} = \frac{1}{a}\sum_{n=0}^{\infty} P_n[\cos(\theta)]\left(\frac{r}{a}\right)^n. \tag{9.3.159}$$

From Equation 9.3.155,

$$\sum_{n=0}^{\infty}\left[\frac{q}{a}\left(\frac{r_0}{a}\right)^n + B_n\right]P_n[\cos(\theta)] = 0. \tag{9.3.160}$$

We can only satisfy Equation 9.3.160 if the square-bracketed term vanishes identically so that

$$B_n = -\frac{q}{a}\left(\frac{r_0}{a}\right)^n. \tag{9.3.161}$$

On substituting Equation 9.3.161 back into Equation 9.3.154,

$$v(r,\theta) = -\frac{qr_0}{ra}\sum_{n=0}^{\infty}\left(\frac{r_0^2}{ar}\right)^n P_n[\cos(\theta)]. \tag{9.3.162}$$

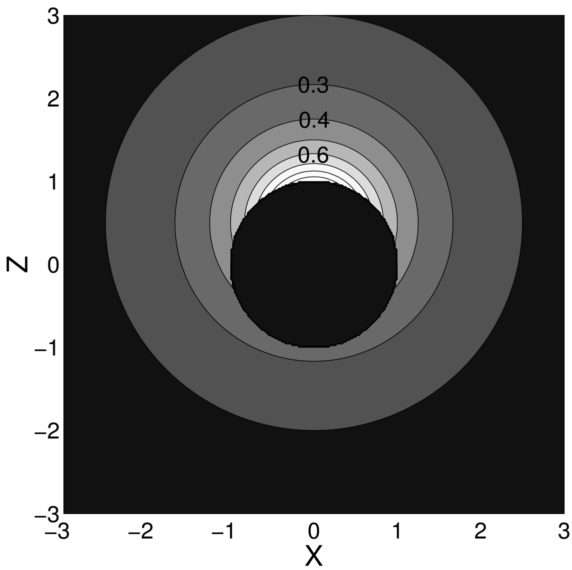


Figure 9.3.11: Electrostatic potential outside of a grounded conducting sphere in the presence of a point charge located at $a/r_0 = 2$. Contours are in units of $-q/r_0$.

The physical interpretation of Equation 9.3.162 is as follows: Consider a point, such as a' (see Figure 9.3.10) on the z-axis. If $r > a'$, the Legendre expansion of $1/s'$ is

$$\frac{1}{s'} = \frac{1}{r} \sum_{n=0}^{\infty} P_n[\cos(\theta)] \left(\frac{a'}{r}\right)^n, \quad r > a'. \tag{9.3.163}$$

Using Equation 9.3.163, we can rewrite it as

$$v(r,\theta) = -\frac{qr_0}{as'}, \tag{9.3.164}$$

if we set $a' = r_0^2/a$. Our final result is then

$$u(r,\theta) = \frac{q}{s} - \frac{q'}{s'}, \tag{9.3.165}$$

provided that q' equals $r_0 q/a$. In other words, when we place a grounded conducting sphere near a point charge $+q$, it changes the potential in the same manner as would a point charge of the opposite sign and magnitude $q' = r_0 q/a$, placed at the point $a' = r_0^2/a$. The charge q' is the *image* of q.

Figure 9.3.11 illustrates the solution, Equation 9.3.162, and was created using the MATLAB script

```
clear
a_over_r0 = 2;
% set up x-z array
dx = 0.02; x = -3:dx:3; dz = 0.02; z = -3:dz:3;
u = 1000 * zeros(length(x),length(z));
X = x' * ones(1,length(z)); Z = ones(length(x),1) * z;
% compute r and theta
```

```
rr = sqrt(X .* X + Z .* Z);
theta = atan2(X,Z);
% find the potential
r_over_aprime = a_over_r0 * rr;
s = 1 + r_over_aprime .* r_over_aprime ...
    - 2 * r_over_aprime .* cos(theta);
for j = 1:length(z); for i = 1:length(x);
  if rr(i,j) >= 1; u(i,j) = 1 ./ sqrt(s(i,j)); end;
end; end
% plot the solution
[cs,h] = contourf(X,Z,u); colormap(hot); brighten(hot,0.5);
axis square; clabel(cs,h,'manual','Fontsize',16);
xlabel('X','Fontsize',20); ylabel('Z','Fontsize',20);
```

Because the charge is located directly above the sphere, the electrostatic potential for any fixed r is largest at the point $\theta = 0$ and weakest at $\theta = \pi$. □

• Example 9.3.10: Poisson's integral formula

In this example we find the solution to Laplace's equation within a unit disc. The problem can be posed as

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} = 0, \quad 0 \le r < 1, \quad 0 \le \varphi \le 2\pi, \tag{9.3.166}$$

with the boundary condition $u(1, \varphi) = f(\varphi)$.

We begin by assuming the separable solution $u(r, \varphi) = R(r)\Phi(\varphi)$ so that

$$\frac{r^2 R'' + rR'}{R} = -\frac{\Phi''}{\Phi} = k^2. \tag{9.3.167}$$

The solution to $\Phi'' + k^2\Phi = 0$ is

$$\Phi(\varphi) = A\cos(k\varphi) + B\sin(k\varphi). \tag{9.3.168}$$

The solution to $R(r)$ is

$$R(r) = Cr^k + Dr^{-k}. \tag{9.3.169}$$

Because the solution must be bounded for all r and periodic in φ, we must take $D = 0$ and $k = n$, where $n = 0, 1, 2, 3, \ldots$. Then, the most general solution is

$$u(r, \varphi) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} \left[a_n \cos(n\varphi) + b_n \sin(n\varphi)\right] r^n, \tag{9.3.170}$$

where a_n and b_n are chosen to satisfy

$$u(1, \varphi) = f(\varphi) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(n\varphi) + b_n \sin(n\varphi). \tag{9.3.171}$$

Because

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos(n\theta)\,d\theta, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin(n\theta)\,d\theta, \tag{9.3.172}$$

we may write $u(r,\varphi)$ as

$$u(r,\varphi) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\left\{\tfrac{1}{2} + \sum_{n=1}^{\infty} r^n \cos[n(\theta-\varphi)]\right\} d\theta. \tag{9.3.173}$$

If we let $\alpha = \theta - \varphi$, and $z = r[\cos(\alpha) + i\sin(\alpha)]$, then

$$\sum_{n=0}^{\infty} r^n \cos(n\alpha) = \Re\left(\sum_{n=0}^{\infty} z^n\right) = \Re\left(\frac{1}{1-z}\right) = \Re\left[\frac{1}{1 - r\cos(\alpha) - ir\sin(\alpha)}\right] \tag{9.3.174}$$

$$= \Re\left[\frac{1 - r\cos(\alpha) + ir\sin(\alpha)}{1 - 2r\cos(\alpha) + r^2}\right] \tag{9.3.175}$$

for all r such that $|r| < 1$. Consequently,

$$\sum_{n=0}^{\infty} r^n \cos(n\alpha) = \frac{1 - r\cos(\alpha)}{1 - 2r\cos(\alpha) + r^2} \tag{9.3.176}$$

$$\frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos(n\alpha) = \frac{1 - r\cos(\alpha)}{1 - 2r\cos(\alpha) + r^2} - \frac{1}{2} \tag{9.3.177}$$

$$= \frac{1}{2}\frac{1 - r^2}{1 - 2r\cos(\alpha) + r^2}. \tag{9.3.178}$$

Substituting Equation 9.3.178 into Equation 9.3.173, we finally have that

$$u(r,\varphi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\frac{1 - r^2}{1 - 2r\cos(\theta-\varphi) + r^2}\, d\theta. \tag{9.3.179}$$

This solution to Laplace's equation within the unit circle is referred to as *Poisson's integral formula.*[15]

Problems

Rectangular Coordinates

Solve Laplace's equation over the rectangular region $0 < x < a$, $0 < y < b$ with the following boundary conditions. Illustrate your solution using MATLAB.

1. $u(x,0) = u(x,b) = u(a,y) = 0, u(0,y) = 1$

2. $u(x,0) = u(0,y) = u(a,y) = 0, u(x,b) = x$

3. $u(x,0) = u(0,y) = u(a,y) = 0, u(x,b) = x - a$

4. $u(x,0) = u(0,y) = u(a,y) = 0,$

[15] Poisson, S. D., 1820: Mémoire sur la manière d'exprimer les fonctions par des séries de quantités périodiques, et sur l'usage de cette transformation dans la résolution de différens problèmes. *J. École Polytech.*, **18**, 417–489.

$$u(x,b) = \begin{cases} 2x/a, & 0 < x < a/2, \\ 2(a-x)/a, & a/2 < x < a. \end{cases}$$

5. $u_x(0,y) = u(a,y) = u(x,0) = 0, u(x,b) = 1$

6. $u_y(x,0) = u(x,b) = u(a,y) = 0, u(0,y) = 1$

7. $u_y(x,0) = u_y(x,b) = 0, u(0,y) = u(a,y) = 1$

8. $u_x(a,y) = u_y(x,b) = 0, u(0,y) = u(x,0) = 1$

9. $u_y(x,0) = u(x,b) = 0, u(0,y) = u(a,y) = 1$

10. $u(a,y) = u(x,b) = 0, u(0,y) = u(x,0) = 1$

11. $u_x(0,y) = 0, u(a,y) = u(x,0) = u(x,b) = 1$

12. $u_x(0,y) = u_x(a,y) = 0, u(x,b) = u_1$,

$$u(x,0) = \begin{cases} f(x), & 0 < x < \alpha, \\ 0, & \alpha < x < a. \end{cases}$$

13. Variations in the earth's surface temperature can arise as a result of topographic undulations and the altitude dependence of the atmospheric temperature. These variations, in turn, affect the temperature within the solid earth. To show this, solve Laplace's equation with the surface boundary condition that

$$u(x,0) = T_0 + \Delta T \cos(2\pi x/\lambda),$$

where λ is the wavelength of the spatial temperature variation. What must be the condition on $u(x,y)$ as we go towards the center of the earth (i.e., $y \to \infty$)?

14. Tóth[16] generalized his earlier analysis of groundwater in an aquifer when the water table follows the topography. Find the groundwater potential if it varies as $u(x,z_0) = g[z_0 + cx + a\sin(bx)]$ at the surface $y = z_0$, while $u_x(0,y) = u_x(L,y) = u_y(x,0) = 0$, where g is the acceleration due to gravity. Assume that $bL \neq n\pi$, where $n = 1,2,3,\ldots$.

15. During his study of fluid flow within a packed bed, Grossman[17] solved

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < 1, \quad 0 < y < L,$$

subject to the boundary conditions

$$u(x,0) = L, \quad u(x,L) = 0, \qquad 0 < x < 1,$$

and

$$u_x(0,y) = 0, \quad u_x(1,y) = -\gamma, \qquad 0 < y < L.$$

[16] Tóth, J. A., 1963: A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.*, **68**, 4795–4812.

[17] Grossman, G., 1975: Stresses and friction forces in moving packed beds. *AICHE J.*, **21**, 720–730.

What should he have found? Hint: Introduce $u(x, y) = L - y + \gamma v(x, y)$.

Cylindrical Coordinates in Finite Domains

16. Solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < a, \quad -L < z < L,$$

with

$$u(a, z) = 0, \quad \text{and} \quad \frac{\partial u(r, -L)}{\partial z} = \frac{\partial u(r, L)}{\partial z} = 1.$$

17. During their study of the role that diffusion plays in equalizing gas concentrations within that portion of the lung that is connnected to terminal bronchioles, Chang et al.[18] solved Laplace's equation in cylindrical coordinates

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < b, \quad -L < z < L,$$

subject to the boundary conditions that

$$\lim_{r \to 0} |u(r, z)| < \infty, \qquad \frac{\partial u(b, z)}{\partial r} = 0, \qquad -L < z < L,$$

and

$$\frac{\partial u(r, -L)}{\partial z} = \frac{\partial u(r, L)}{\partial z} = \begin{cases} A, & 0 \le r < a, \\ 0, & a < r < b. \end{cases}$$

What should they have found?

18. Solve[19]

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < b, \quad 0 < z < L,$$

with the boundary conditions

$$\lim_{r \to 0} |u(r, z)| < \infty, \qquad \frac{\partial u(b, z)}{\partial r} = 0, \qquad 0 \le z \le L,$$

$$u(r, L) = A, \qquad 0 \le r \le b,$$

and

$$\frac{\partial u(r, 0)}{\partial z} = \begin{cases} B, & 0 \le r < a, \\ 0, & a < r < b. \end{cases}$$

19. Solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < a, \quad 0 < z < h,$$

---

[18] Chang, D. B., S. M. Lewis, and A. C. Young, 1976: A theoretical discussion of diffusion and convection in the lung. *Math. Biosci.*, **29**, 331–349.

[19] See Keller, K. H., and T. R. Stein, 1967: A two-dimensional analysis of porous membrane transfer. *Math. Biosci.*, **1**, 421–437.

with

$$\frac{\partial u(a,z)}{\partial r} = u(r,h) = 0$$

and

$$\frac{\partial u(r,0)}{\partial z} = \begin{cases} 1, & 0 \le r < r_0, \\ 0, & r_0 < r < a. \end{cases}$$

20. Solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < 1, \quad 0 < z < d,$$

with

$$\frac{\partial u(1,z)}{\partial r} = \frac{\partial u(r,0)}{\partial z} = 0,$$

and

$$u(r,d) = \begin{cases} -1, & 0 \le r < a, \quad b < r < 1, \\ 1/(b^2 - a^2) - 1, & a < r < b. \end{cases}$$

21. Solve[20]

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < 1, \quad 0 < z < 1.$$

Take for the boundary conditions either (a)

$$\lim_{r\to 0} |u(r,z)| < \infty, \quad u(1,z) = -1, \qquad 0 < z < 1,$$

and

$$u_z(r,0) = u(r,1) = 0, \qquad 0 < r < 1;$$

or (b)

$$\lim_{r\to 0} |u(r,z)| < \infty, \quad u(1,z) = 0, \qquad 0 < z < 1,$$

and

$$u_z(r,0) = 0, \quad u(r,1) = r, \qquad 0 < r < 1.$$

22. Solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < a, \quad 0 < z < h,$$

with

$$\lim_{r\to 0} |u(r,z)| < \infty, \quad u(a,z) = 0, \qquad 0 < z < h,$$

and

$$u(r,0) = 0, \quad u_z(r,h) = Ar, \qquad 0 \le r < a.$$

23. Solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < a, \quad 0 < z < 1,$$

[20] See Muite, B. K., 2004: The flow in a cylindrical container with a rotating end wall at small but finite Reynolds number. *Phys. Fluids*, **16**, 3614–3626.

with

$$\lim_{r\to 0}|u(r,z)| < \infty, \quad u(a,z) = z, \qquad 0 < z < 1,$$

and

$$u(r,0) = u(r,1) = 0, \qquad 0 \le r < a.$$

24. Solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < a, \quad 0 < z < h,$$

with

$$\lim_{r\to 0}|u(r,z)| < \infty, \quad \frac{\partial u(a,z)}{\partial r} = 0, \qquad 0 < z < h,$$

and

$$u(r,0) = 0, \quad \frac{\partial u(r,h)}{\partial z} = r, \qquad 0 \le r < a.$$

25. Solve

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} - \frac{u}{r^2} + \frac{\partial^2 u}{\partial z^2} = 0, \qquad 0 \le r < 1, \quad -a < z < a,$$

with the boundary conditions

$$\lim_{r\to 0}|u(r,z)| < \infty, \qquad \frac{\partial u(1,z)}{\partial r} = u(1,z), \qquad -a < z < a,$$

and

$$-\frac{\partial u(r,-a)}{\partial z} = \frac{\partial u(r,a)}{\partial z} = r, \qquad 0 \le r < 1.$$

26. Solve[21]

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} - u = 0, \qquad 0 \le r < 1, \quad 0 < z < L,$$

subject to the boundary conditions that

$$\lim_{r\to 0}|u(r,z)| < \infty, \quad u_r(1,z) = -h\,u(1,z), \qquad 0 < z < L,$$

and

$$u(r,0) = u_0, \quad u(r,L) = 0, \qquad 0 \le r < 1,$$

using the Fourier-Bessel series

$$u(r,z) = \sum_{n=1}^{\infty} A_n Z_n(z) J_0(k_n r)$$

where k_n is the nth root of $k\,J_0'(k) + h\,J_0(k) = h\,J_0(k) - k\,J_1(k) = 0$.

[21] See Stripp, K. F., and A. R. Moore, 1955: The effects of junction shape and surface recombination on transistor current gain – Part II. *Proc. IRE*, **43**, 856–866.

27. Solve[22] Laplace's equation in cylindrical coordinates

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \qquad 0 \le r < a, \quad 0 < z < L,$$

subject to the boundary conditions that

$$\lim_{r\to 0} |u(r,z)| < \infty, \quad -Du_r(a,z) = Ku(a,z), \qquad 0 < z < L,$$

and

$$u(r,0) = u_0, \quad u_z(r,L) = 0, \qquad 0 \le r < a.$$

28. Solve[23] the partial differential equation

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = a^2 u, \qquad 0 \le r < 1, \quad 0 < z < 1,$$

subject to the boundary conditions that

$$\lim_{r\to 0} |u(r,z)| < \infty, \quad u(1,z) = 1, \qquad 0 < z < 1,$$

and

$$u(r,0) = 1, \quad u(r,1) = 1, \qquad 0 \le r < 1.$$

Hint: Break the problem into three parts: $u(r,z) = u_1(r,z) + u_2(r,z) + u_3(r,z)$, where

$$\frac{\partial^2 u_1}{\partial r^2} + \frac{1}{r}\frac{\partial u_1}{\partial r} + \frac{\partial^2 u_1}{\partial z^2} = a^2 u_1, \qquad 0 \le r < 1, \quad 0 < z < 1,$$

subject to the boundary conditions that

$$\lim_{r\to 0} |u_1(r,z)| < \infty, \quad u_1(1,z) = 1, \qquad 0 < z < 1,$$

and

$$u_1(r,0) = 0, \quad u_1(r,1) = 0, \qquad 0 \le r < 1;$$

$$\frac{\partial^2 u_2}{\partial r^2} + \frac{1}{r}\frac{\partial u_2}{\partial r} + \frac{\partial^2 u_2}{\partial z^2} = a^2 u_2, \qquad 0 \le r < 1, \quad 0 < z < 1,$$

subject to the boundary conditions that

$$\lim_{r\to 0} |u_2(r,z)| < \infty, \quad u_2(1,z) = 0, \qquad 0 < z < 1,$$

[22] See Bischoff, K. B., 1966: Transverse diffusion in catalyst pores. *Indust. Engng. Chem. Fund.*, **5**, 135–136.

[23] See Gunn, D. J., 1967: Diffusion and chemical reaction in catalysis and absorption. *Chem. Engng. Sci.*, **22**, 1439–1455; Ho, T. C., and G. C. Hsiao, 1977: Estimation of the effectiveness factor for a cylindrical catalyst support: A singular perturbation approach. *Chem. Engng. Sci.*, **32**, 63–66.

and

$$u_2(r,0) = 1, \quad u_2(r,1) = 0, \qquad 0 \le r < 1;$$

and

$$\frac{\partial^2 u_3}{\partial r^2} + \frac{1}{r}\frac{\partial u_3}{\partial r} + \frac{\partial^2 u_3}{\partial z^2} = a^2 u_3, \qquad 0 \le r < 1, \quad 0 < z < 1,$$

subject to the boundary conditions that

$$\lim_{r\to 0} |u_3(r,z)| < \infty, \quad u_3(1,z) = 0, \qquad 0 < z < 1,$$

and

$$u_3(r,0) = 0, \quad u_3(r,1) = 0, \qquad 0 \le r < 1.$$

Cylindrical Coordinates on Infinite and Semi-Infinite Domains

29. Solve[24] Laplace's equation in cylindrical coordinates

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \qquad 0 \le r < a, \quad -\infty < z < \infty,$$

subject to the boundary conditions that

$$\lim_{r\to 0} |u(r,z)| < \infty, \qquad u(a,z) = \begin{cases} -V, & |z| < d/2, \\ 0, & |z| > d/2, \end{cases}$$

and

$$\lim_{|z|\to\infty} u(r,z) \to 0, \qquad 0 \le r < a.$$

Hint: Show that the solution can be written

$$u(r,z) = -V + \sum_{n=1}^{\infty} A_n \cosh\left(\frac{k_n z}{a}\right) J_0\left(\frac{k_n r}{a}\right), \qquad |z| < d/2,$$

and

$$u(r,z) = \sum_{n=1}^{\infty} B_n \exp\left(-\frac{k_n |z|}{a}\right) J_0\left(\frac{k_n r}{a}\right), \qquad |z| > d/2$$

with the additional conditions that

$$u(r, d^-/2) = u(r, d^+/2) \text{ and } u_z(r, d^-/2) = u_z(r, d^+/2), \quad 0 \le r < a,$$

where k_n is the nth root of $J_0(k) = 0$. Then find A_n and B_n.

30. Solve Laplace's equation in cylindrical coordinates

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \qquad 0 \le r < b, \quad 0 < z < \infty,$$

---

[24] See Striffler, C. D., C. A. Kapetanakos, and R. C. Davidson, 1975: Equilibrium properties of a rotating nonneutral E layer in a coupled magnetic field. *Phys. Fluids*, **18**, 1374–1382.

subject to the boundary conditions that

$$\lim_{r\to 0}|u(r,z)| < \infty, \qquad \frac{\partial u(b,z)}{\partial r} = 0, \qquad 0 < z < \infty,$$

and

$$\lim_{z\to\infty}|u(r,z)| < \infty, \qquad u(r,0) = \begin{cases} A, & 0 \le r < a, \\ 0, & a < r < b. \end{cases}$$

31. Solve[25]

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} - \frac{\partial u}{\partial z} = 0, \quad 0 \le r < 1, \quad 0 < z < \infty,$$

with the boundary conditions

$$\lim_{r\to 0}|u(r,z)| < \infty, \qquad \frac{\partial u(1,z)}{\partial r} = -Bu(1,z), \qquad 0 < z,$$

and

$$u(r,0) = 1, \qquad \lim_{z\to\infty}|u(r,z)| < \infty, \qquad 0 \le r < 1,$$

where B is a constant.

32. Solve[26]

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} - \frac{1}{H}\frac{\partial u}{\partial z} = 0, \quad 0 \le r < b, \quad 0 < z < \infty,$$

with the boundary conditions

$$\lim_{r\to 0}|u(r,z)| < \infty, \qquad \frac{\partial u(b,z)}{\partial r} = -hu(b,z), \qquad 0 < z,$$

$$\lim_{z\to\infty}|u(r,z)| < \infty, \qquad 0 \le r < b,$$

and

$$\frac{u(r,0)}{H} - u_z(r,0) = \begin{cases} Q, & 0 \le r < a, \\ 0, & a \le r < b, \end{cases}$$

where $b > a$.

Mixed Boundary-Value Problems

33. Solve[27]

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial z^2} = 0, \quad 0 \le r < \infty, \quad 0 < z < \infty,$$

---

[25] See Kern, J., and J. O. Hansen, 1976: Transient heat conduction in cylindrical systems with an axially moving boundary. *Int. J. Heat Mass Transfer*, **19**, 707–714.

[26] See Smirnova, E. V., and I. A. Krinberg, 1970: Spatial distribution of the atoms of an impurity element in an arc discharge. I. *J. Appl. Spectroscopy*, **13**, 859–864.

[27] See Fleischmann, M., and S. Pons, 1987: The behavior of microdisk and microring electrodes. *J. Electroanal. Chem.*, **222**, 107–115.

subject to the boundary conditions

$$\lim_{r\to 0} |u(r,z)| < \infty, \qquad \lim_{r\to\infty} |u(r,z)| < \infty, \qquad 0 < z < \infty,$$

$$\lim_{z\to\infty} u(r,z) \to u_\infty, \qquad 0 \le r < \infty,$$

and the mixed boundary condition

$$\begin{cases} u(r,0) = u_\infty - \Delta u, & 0 \le r < a, \\ u_z(r,0) = 0, & a \le r < \infty, \end{cases}$$

where u_∞ and Δu are constants.

Step 1: Show that

$$u(r,z) = u_\infty - \int_0^\infty A(k) e^{-kz} J_0(kr)\, dk$$

satisfies the partial differential equation and the boundary conditions as $r \to 0$, $r \to \infty$, and $z \to \infty$.

Step 2: Show that

$$\int_0^\infty kA(k) J_0(kr)\, dk = 0, \qquad a < r < \infty.$$

Step 3: Using the relationship[28]

$$\int_0^\infty \sin(ka) J_0(kr)\, dk = \begin{cases} (a^2 - r^2)^{-\frac{1}{2}}, & r < a, \\ 0, & r > a. \end{cases}$$

show that $kA(k) = C \sin(ka)$.

Step 4: Using the relationship[29]

$$\int_0^\infty \sin(ka) J_0(kr)\, \frac{dk}{k} = \begin{cases} \pi/2, & r \le a, \\ \sin^{-1}(a/r), & r \ge a, \end{cases}$$

show that

$$u(r,z) = u_\infty - \frac{2\Delta u}{\pi} \int_0^\infty e^{-kz} \sin(ka) J_0(kr)\, \frac{dk}{k}.$$

34. Solve Laplace's equation[30]

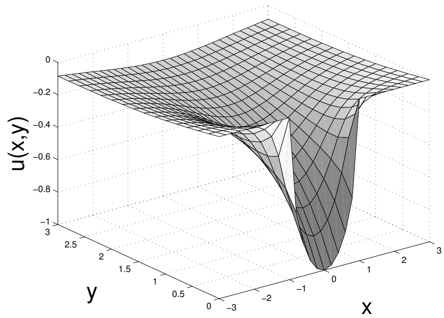
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \qquad -\infty < x < \infty, \quad 0 < y < \infty,$$

---

[28] Gradshteyn and Ryzhik, op. cit., Section 6.671, Formula 7.

[29] Ibid., Section 6.693, Formula 1 with $\nu = 0$.

[30] See Yang, F.-Q., and J. C. M. Li, 1995: Impression and diffusion creep of anisotropic media. *J. Appl. Phys.*, **77**, 110–117. See also Shindo, Y., H. Tamura, and Y. Atobe, 1990: Transient singular stresses of a finite crack in an elastic conductor under electromagnetic force (in Japanese). *Nihon Kikai Gakkai Rombunshu (Trans. Japan Soc. Mech. Engrs.), Ser. A*, **56**, 278–282.



Problem 34

with the boundary conditions

$$\lim_{|x|\to\infty} u(x,y) \to 0, \qquad 0 < y < \infty,$$

$$\lim_{y\to\infty} u(x,y) \to 0, \qquad -\infty < x < \infty,$$

and the mixed boundary condition

$$\begin{cases} u_y(x,0) = 1, & 0 \le |x| < 1, \\ u(x,0) = 0, & 1 < |x| < \infty. \end{cases} \tag{1}$$

This problem would arise in finding the electrostatic potential inside the half-space $y > 0$ when the boundary $x = 0$ is grounded for $|x| > 1$ and the vertical electric field equals one for $|x| < 1$.

Step 1: Using separation of variables or transform methods, show that the general solution to the problem is

$$u(x,y) = \int_0^\infty A(k)e^{-ky}\cos(kx)\,dk.$$

Step 2: Using boundary condition (1), show that $A(k)$ satisfies the dual integral equations

$$\int_0^\infty kA(k)\cos(kx)\,dk = -\frac{\pi}{2}, \qquad 0 \le |x| < 1,$$

and

$$\int_0^\infty A(k)\cos(kx)\,dk = 0, \qquad 1 < |x| < \infty.$$

Step 3: Using integral tables,[31] show that

$$A(k) = -\frac{\pi J_1(k)}{2k}$$

satisfies both integral equations given in Step 2.

Step 4: Show that the solution to this problem is

$$u(x,y) = -\int_0^\infty J_1(k)e^{-ky}\cos(kx)\frac{dk}{k}.$$

In particular, verify that $u(x,0) = -\sqrt{1-x^2}$ if $|x| < 1$.

[31] Gradshteyn and Ryzhik, op. cit., Section 6.671, Formula 2 and Section 6.693, Formula 2 with $\nu = 1$.

Spherical Coordinates

35. Find the steady-state temperature within a sphere of radius a if the temperature along its surface is maintained at the temperature $u(a, \theta) = 100[\cos(\theta) - \cos^5(\theta)]$.

36. Find the steady-state temperature within a sphere if the upper half of the exterior surface at radius a is maintained at the temperature 100 while the lower half is maintained at the temperature 0.

37. The surface of a sphere of radius a has a temperature of zero everywhere except in a spherical cap at the north pole (defined by the cone $\theta = \alpha$), where it equals T_0. Find the steady-state temperature within the sphere.

Poisson's Integral Formula

38. Using the relationship

$$\int_0^{2\pi} \frac{d\varphi}{1 - b\cos(\varphi)} = \frac{2\pi}{\sqrt{1 - b^2}}, \qquad |b| < 1$$

and Poisson's integral formula, find the solution to Laplace's equation within a unit disc if $u(1, \varphi) = f(\varphi) = T_0$, a constant.

9.4 POISSON'S EQUATION ON A RECTANGLE

Poisson's equation[32] is Laplace's equation with a source term:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y). \tag{9.4.1}$$

It arises in such diverse areas as groundwater flow, electromagnetism, and potential theory. Let us solve it if $u(0, y) = u(a, y) = u(x, 0) = u(x, b) = 0$.

We begin by solving a similar partial differential equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \lambda u, \quad 0 < x < a, \quad 0 < y < b, \tag{9.4.2}$$

by separation of variables. If $u(x, y) = X(x)Y(y)$, then

$$\frac{X''}{X} + \frac{Y''}{Y} = \lambda. \tag{9.4.3}$$

Because we must satisfy the boundary conditions that $X(0) = X(a) = Y(0) = Y(b) = 0$, we have the following eigenfunction solutions:

$$X_n(x) = \sin\left(\frac{n\pi x}{a}\right), \qquad Y_m(x) = \sin\left(\frac{m\pi y}{b}\right) \tag{9.4.4}$$

[32] Poisson, S. D., 1813: Remarques sur une équation qui se présente dans la théorie des attractions des sphéroïdes. *Nouv. Bull. Soc. Philomath. Paris*, **3**, 388–392.

Siméon-Denis Poisson (1781–1840) was a product as well as a member of the French scientific establishment of his day. Educated at the École Polytechnique, he devoted his life to teaching, both in the classroom and with administrative duties, and to scientific research. Poisson's equation dates from 1813 when Poisson sought to extend Laplace's work on gravitational attraction. (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

with $\lambda_{nm} = -n^2\pi^2/a^2 - m^2\pi^2/b^2$; otherwise, we would only have trivial solutions. The corresponding particular solutions are

$$u_{nm} = A_{nm} \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right), \tag{9.4.5}$$

where $n = 1, 2, 3, \ldots$, and $m = 1, 2, 3, \ldots$.

For a fixed y, we can expand $f(x, y)$ in the half-range Fourier sine series

$$f(x, y) = \sum_{n=1}^{\infty} A_n(y) \sin\left(\frac{n\pi x}{a}\right), \tag{9.4.6}$$

where

$$A_n(y) = \frac{2}{a} \int_0^a f(x, y) \sin\left(\frac{n\pi x}{a}\right) \, dx. \tag{9.4.7}$$

However, we can also expand $A_n(y)$ in a half-range Fourier sine series

$$A_n(y) = \sum_{m=1}^{\infty} a_{nm} \sin\left(\frac{m\pi y}{b}\right), \tag{9.4.8}$$

where

$$a_{nm} = \frac{2}{b}\int_0^b A_n(y)\sin\left(\frac{m\pi y}{b}\right)\,dy = \frac{4}{ab}\int_0^b\int_0^a f(x,y)\sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right)\,dx\,dy, \tag{9.4.9}$$

and

$$f(x,y) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{nm}\sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right). \tag{9.4.10}$$

In other words, we re-expressed $f(x,y)$ in terms of a *double Fourier series*.

Because Equation 9.4.2 must hold for each particular solution,

$$\frac{\partial^2 u_{nm}}{\partial x^2} + \frac{\partial^2 u_{nm}}{\partial y^2} = \lambda_{nm}u_{nm} = a_{nm}\sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right), \tag{9.4.11}$$

if we now associate Equation 9.4.1 with Equation 9.4.2. Therefore, the solution to Poisson's equation on a rectangle where the boundaries are held at zero is the double Fourier series

$$u(x,y) = -\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{a_{nm}}{n^2\pi^2/a^2 + m^2\pi^2/b^2}\sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right). \tag{9.4.12}$$

Problems

1. The equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -\frac{R}{T}, \quad |x| < a, \quad |y| < b$$

describes the hydraulic potential (elevation of the water table) $u(x,y)$ within a rectangular island on which a recharging well is located at $(0,0)$. Here R is the rate of recharging and T is the product of the hydraulic conductivity and aquifer thickness. If the water table is at sea level around the island so that $u(-a,y) = u(a,y) = u(x,-b) = u(x,b) = 0$, find $u(x,y)$ everywhere in the island. Hint: Use symmetry and redo the above analysis with the boundary conditions: $u_x(0,y) = u(a,y) = u_y(x,0) = u(x,b) = 0$.

2. Let us apply the same approach that we used to find the solution of Poisson's equation on a rectangle to solve the axisymmetric Poisson equation inside a circular cylinder

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial z^2} = f(r,z), \quad 0 \le r < a, \quad |z| < b,$$

subject to the boundary conditions

$$\lim_{r\to 0}|u(r,z)| < \infty, \qquad u(a,z) = 0, \qquad |z| < b,$$

and

$$u(r,-b) = u(r,b) = 0, \qquad 0 \le r < a.$$

Step 1: Replace the original problem with

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial z^2} = \lambda u, \quad 0 \le r < a, \quad |z| < b,$$

subject to the same boundary conditions. Use separation of variables to show that the solution to this new problem is

$$u_{nm}(r,z) = A_{nm} J_0\left(k_n \frac{r}{a}\right) \cos\left[\frac{\left(m+\frac{1}{2}\right)\pi z}{b}\right],$$

where k_n is the nth zero of $J_0(k) = 0$, $n = 1, 2, 3, \ldots$, and $m = 0, 1, 2, \ldots$.

Step 2: Show that $f(r,z)$ can be expressed as

$$f(r,z) = \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} a_{nm} J_0\left(k_n \frac{r}{a}\right) \cos\left[\frac{\left(m+\frac{1}{2}\right)\pi z}{b}\right],$$

where

$$a_{nm} = \frac{2}{a^2 b J_1^2(k_n)} \int_{-b}^{b} \int_0^a f(r,z) J_0\left(k_n \frac{r}{a}\right) \cos\left[\frac{\left(m+\frac{1}{2}\right)\pi z}{b}\right] r\, dr\, dz.$$

Step 3: Show that the general solution is

$$u(r,z) = -\sum_{n=1}^{\infty} \sum_{m=0}^{\infty} a_{nm} \frac{J_0(k_n r/a) \cos\left[\left(m+\frac{1}{2}\right)\pi z/b\right]}{(k_n/a)^2 + \left[\left(m+\frac{1}{2}\right)\pi/b\right]^2}.$$

9.5 NUMERICAL SOLUTION OF LAPLACE'S EQUATION

As in the case of the heat and wave equations, numerical methods can be used to solve elliptic partial differential equations when analytic techniques fail or are too cumbersome. They are also employed when the domain differs from simple geometries.

The numerical analysis of an elliptic partial differential equation begins by replacing the continuous partial derivatives by finite-difference formulas. Employing centered differencing,

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{(\Delta x)^2} + O[(\Delta x)^2], \qquad \textbf{(9.5.1)}$$

and

$$\frac{\partial^2 u}{\partial y^2} = \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{(\Delta y)^2} + O[(\Delta y)^2], \qquad \textbf{(9.5.2)}$$

where $u_{m,n}$ denotes the solution value at the grid point m,n. If $\Delta x = \Delta y$, Laplace's equation becomes the difference equation

$$u_{m+1,n} + u_{m-1,n} + u_{m,n+1} + u_{m,n-1} - 4u_{m,n} = 0. \qquad \textbf{(9.5.3)}$$

Thus, we must now solve a set of simultaneous linear equations that yield the value of the solution at each grid point.

The solution of Equation 9.5.3 is best done using techniques developed by algebraists. Later on, in Chapter 3, we will show that a very popular method for directly solving systems of linear equations is Gaussian elimination. However, for many grids at a reasonable resolution, the number of equations is generally in the tens of thousands. Because most of the coefficients in the equations are zero, Gaussian elimination is unsuitable, both from the point of view of computational expense and accuracy. For this reason alternative methods have been developed that generally use successive corrections or iterations. The most common of these point iterative methods are the Jacobi method, unextrapolated Liebmann or Gauss-Seidel method, and extrapolated Liebmann or successive over-relaxation (SOR). None of these approaches is completely satisfactory because of questions involving convergence and efficiency. Because of its simplicity we will focus on the Gauss-Seidel method.

We may illustrate the Gauss-Seidel method by considering the system:

$$10x + y + z = 39, \tag{9.5.4}$$

$$2x + 10y + z = 51, \tag{9.5.5}$$

and

$$2x + 2y + 10z = 64. \tag{9.5.6}$$

An important aspect of this system is the dominance of the coefficient of x in the first equation of the set and that the coefficients of y and z are dominant in the second and third equations, respectively.

The Gauss-Seidel method may be outlined as follows:

- Assign an initial value for each unknown variable. If possible, make a good first guess. If not, any arbitrarily selected values may be chosen. The initial value will not affect the convergence but will affect the number of iterations until convergence.

- Starting with Equation 9.5.4, solve that equation for a new value of the unknown which has the largest coefficient in that equation, using the assumed values for the other unknowns.

- Go to Equation 9.5.5 and employ the same technique used in the previous step to compute the unknown that has the largest coefficient in that equation. Where possible, use the latest values.

- Proceed to the remaining equations, always solving for the unknown having the largest coefficient in the particular equation and always using the *most recently* calculated values for the other unknowns in the equation. When the last equation, Equation 9.5.6, has been solved, you have completed a single iteration.

- Iterate until the value of each unknown does not change within a predetermined value.

Usually a compromise must be struck between the accuracy of the solution and the desired rate of convergence. The more accurate the solution is, the longer it will take for the solution to converge.

To illustrate this method, let us solve our system Equation 9.5.4 through Equation 9.5.6 with the initial guess $x = y = z = 0$. The first iteration yields $x = 3.9$, $y = 4.32$, and $z = 4.756$. The second iteration yields $x = 2.9924$, $y = 4.02592$, and $z = 4.996336$. As can be readily seen, the solution is converging to the correct solution of $x = 3$, $y = 4$, and $z = 5$.

Applying these techniques to Equation 9.5.3,

$$u_{m,n}^{k+1} = \tfrac{1}{4}\left(u_{m+1,n}^{k} + u_{m-1,n}^{k+1} + u_{m,n+1}^{k} + u_{m,n-1}^{k+1}\right), \tag{9.5.7}$$

where we assume that the calculations occur in order of increasing m and n.

• Example 9.5.1

To illustrate the numerical solution of Laplace's equation, let us redo Example 9.3.1 with the boundary condition along $y = H$ simplified to $u(x, H) = 1 + x/L$.

We begin by finite-differencing the boundary conditions. The condition $u_x(0, y) = u_x(L, y) = 0$ leads to $u_{1,n} = u_{-1,n}$ and $u_{M+1,n} = u_{M-1,n}$ if we employ centered differences at $m = 0$ and $m = M$. Substituting these values in Equation 9.5.7, we have the following equations for the left and right boundaries:

$$u_{0,n}^{k+1} = \tfrac{1}{4}\left(2u_{1,n}^{k} + u_{0,n+1}^{k} + u_{0,n-1}^{k+1}\right) \tag{9.5.8}$$

and

$$u_{M,n}^{k+1} = \tfrac{1}{4}\left(2u_{M-1,n}^{k+1} + u_{M,n+1}^{k} + u_{M,n-1}^{k+1}\right). \tag{9.5.9}$$

On the other hand, $u_y(x, 0) = 0$ yields $u_{m,1} = u_{m,-1}$, and

$$u_{m,0}^{k+1} = \tfrac{1}{4}\left(u_{m+1,0}^{k} + u_{m-1,0}^{k+1} + 2u_{m,1}^{k}\right). \tag{9.5.10}$$

At the bottom corners, Equation 9.5.8 through Equation 9.5.10 simplify to

$$u_{0,0}^{k+1} = \tfrac{1}{2}\left(u_{1,0}^{k} + u_{0,1}^{k}\right) \tag{9.5.11}$$

and

$$u_{L,0}^{k+1} = \tfrac{1}{2}\left(u_{L-1,0}^{k+1} + u_{L,1}^{k}\right). \tag{9.5.12}$$

These equations along with Equation 9.5.7 were solved with the Gauss-Seidel method using the MATLAB script

```
clear
dx = 0.1; x = 0:dx:1; M = 1/dx+1; % M = number of x grid points
dy = 0.1; y = 0:dy:1; N = 1/dy+1; % N = number of y grid points
X = x' * ones(1,N); Y = ones(M,1) * y;
u = zeros(M,N); % create initial guess for the solution
% introduce boundary condition along y = H
for m = 1:M; u(m,N) = 1 + x(m); end
% start Gauss-Seidel method for Laplace's equation
for iter = 1:256
% do the interior first
  for n = 2:N-1; for m = 2:M-1;
    u(m,n) = (u(m+1,n)+u(m-1,n)+u(m,n+1)+u(m,n-1)) / 4;
  end; end
% now do the x = 0 and x = L sides
  for n = 2:N-1
    u(1,n) = (2*u( 2 ,n)+u(1,n+1)+u(1,n-1)) / 4;
    u(M,n) = (2*u(M-1,n)+u(M,n+1)+u(M,n-1)) / 4;
  end
% now do the y = 0 side
  for m = 2:M-1
    u(m,1) = (u(m+1,1)+u(m-1,1)+2*u(m,2)) / 4;
```

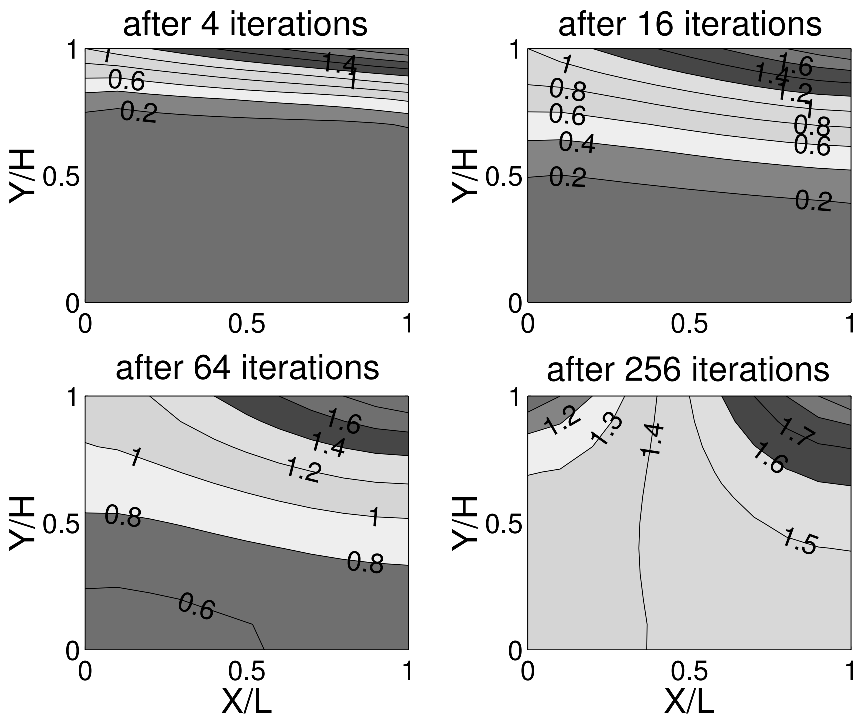


Figure 9.5.1: The solution to Laplace's equation by the Gauss-Seidel method. The boundary conditions are $u_x(0,y) = u_x(L,y) = u_y(x,0) = 0$, and $u(x,H) = 1 + x/L$.

```
  end
% finally do the corners
  u(1,1) = (u(2,1)+u(1,2))/2; u(M,1) = (u(M-1,1)+u(M,2))/2;
% plot the solution
  if (iter == 4) subplot(2,2,1), [cs,h] = contourf(X,Y,u);
    clabel(cs,h,[0.2 0.6 1 1.4],'Fontsize',16)
    axis tight; title('after 4 iterations','Fontsize',20);
    ylabel('Y/H','Fontsize',20); end
  if (iter == 16) subplot(2,2,2), [cs,h] = contourf(X,Y,u);
    clabel(cs,h,'Fontsize',16)
    axis tight; title('after 16 iterations','Fontsize',20);
    ylabel('Y/H','Fontsize',20); end
  if (iter == 64) subplot(2,2,3), [cs,h] = contourf(X,Y,u);
    clabel(cs,h,'Fontsize',16)
    axis tight; title('after 64 iterations','Fontsize',20);
    xlabel('X/L','Fontsize',20); ylabel('Y/H','Fontsize',20);
  end
  if (iter == 256) subplot(2,2,4), [cs,h] = contourf(X,Y,u);
    clabel(cs,h,'Fontsize',16)
    axis tight; title('after 256 iterations','Fontsize',20);
    xlabel('X/L','Fontsize',20); ylabel('Y/H','Fontsize',20);
  end
end
```

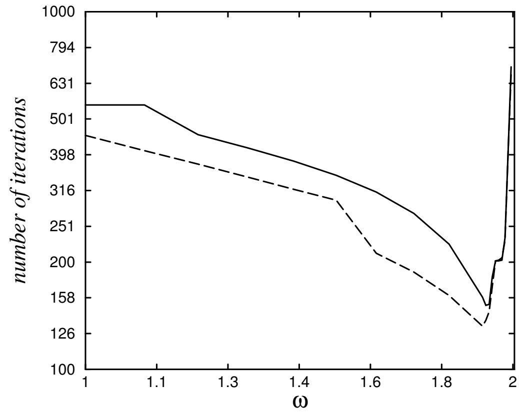


Figure 9.5.2: The number of iterations required so that $|R_{m,n}| \le 10^{-3}$ as a function of ω during the iterative solution of the problem posed in the project. We used $\Delta x = \Delta y = 0.01$, and $L = z_0 = 1$. The iteration count for the boundary conditions stated in Step 1 is given by the solid line while the iteration count for the boundary conditions given in Step 2 is shown by the dotted line. The initial guess equaled zero.

The initial guess everywhere except along the top boundary was zero. In Figure 9.5.1 we illustrate the numerical solution after 4, 16, 64, and 256 iterations, where we have taken 11 grid points in the x and y directions.

Project: Successive Over-Relaxation

The fundamental difficulty with relaxation methods used in solving Laplace's equation is the rate of convergence. Assuming $\Delta x = \Delta y$, the most popular method for accelerating convergence of these techniques is *successive over-relaxation (SOR)*:

$$u_{m,n}^{k+1} = (1-\omega)u_{m,n}^k + \omega R_{m,n},$$

where

$$R_{m,n} = \tfrac{1}{4}\left(u_{m+1,n}^k + u_{m-1,n}^{k+1} + u_{m,n+1}^k + u_{m,n-1}^{k+1}\right).$$

Most numerical methods books dealing with partial differential equations discuss the theoretical reasons behind this technique;[33] the optimum value always lies between one and two. In the present case, a theoretical analysis[34] gives

$$\omega_{\text{opt}} = \frac{4}{2+\sqrt{4-c^2}},$$

where

$$c = \cos\left(\frac{\pi}{N}\right) + \cos\left(\frac{\pi}{M}\right),$$

[33] For example, Young, D. M., 1971: *Iterative Solution of Large Linear Systems*. Academic Press, 570 pp.

[34] Yang, S., and M. K. Gobbert, 2009: The optimal relaxation parameter for the SOR method applied to the Poisson equation in any space dimensions. *Appl. Math. Letters*, **22**, 325–331.

and N and M are the number of mesh divisions on each side of the rectangular domain. Recently Yang and Gobbert[35] generalized the analysis and found the optimal relaxation parameter for the successive-overrelaxation method when it is applied to the Poisson equation in any space dimensions.

Step 1: Write a MATLAB script that uses the Gauss-Seidel method to numerically solve Laplace's equation for $0 \le x \le L$, $0 \le y \le z_0$ with the following boundary conditions: $u(x,0) = 0$, $u(x,z_0) = 1 + x/L$, $u(0,y) = y/z_0$, and $u(L,y) = 2y/z_0$. Because this solution will act as "truth" in this project, you should iterate until the solution does not change.

Step 2: Now redo the calculation using successive over-relaxation. Count the number of iterations until $|R_{m,n}| \le 10^{-3}$ for *all* m and n. Plot the number of iterations as a function of ω. How does the curve change with resolution Δx? How does your answer compare to the theoretical value? See Figure 9.5.2.

Step 3: Redo Steps 1 and 2 with the exception of $u(0,y) = u(L,y) = 0$. How has the convergence rate changed? Can you explain why? How sensitive are your results to the first guess?

9.6 FINITE ELEMENT SOLUTION OF LAPLACE'S EQUATION

In Section 6.6 we showed how the finite element method can be used to solve boundary-value problems. Here we extend this approach to two dimensions. The main difficulty will be the increase in the complexity of the "bookkeeping."

Triangular elements

Our first concern is with the element equations. Essentially there are two types: triangles and quadrilaterals. Triangular elements use the linear polynomial:

$$u(x,y) = a_0 + a_1 x + a_2 y, \tag{9.6.1}$$

where $u(x,y)$ is the dependent variable, the a_i's are coefficients, and x and y are the independent variables. This function must pass through the values of $u(x,y)$ at the triangle's nodes (x_1,y_1), (x_2,y_2), and (x_3,y_3). Therefore,

$$u_1 = a_0 + a_1 x_1 + a_2 y_1, \tag{9.6.2}$$

$$u_2 = a_0 + a_1 x_2 + a_2 y_2, \tag{9.6.3}$$

and

$$u_3 = a_0 + a_1 x_3 + a_2 y_3. \tag{9.6.4}$$

Solving for a_0, a_1 and a_2, we have that

$$a_0 = [u_1(x_2y_3 - x_3y_2) + u_2(x_3y_1 - x_1y_3) + u_3(x_1y_2 - x_2y_1)]/A, \tag{9.6.5}$$

$$a_1 = [u_1(y_2 - y_3) + u_2(y_3 - y_1) + u_3(y_1 - y_2)]/A, \tag{9.6.6}$$

[35] Ibid.

and

$$a_2 = [u_1(x_3 - x_2) + u_2(x_1 - x_3) + u_3(x_2 - x_1)]/A, \qquad \textbf{(9.6.7)}$$

where A is the area of the triangular element

$$A = \tfrac{1}{2}[(x_2y_3 - x_3y_2) + (x_3y_1 - x_1y_3) + (x_1y_2 - x_2y_1)]. \qquad \textbf{(9.6.8)}$$

To avoid a small A you should avoid elements with narrow geometries on the finite mesh. Equation 9.6.5 through Equation 9.6.7 can be substituted into Equation 9.6.1. Collecting terms, the result can be expressed as

$$u(x, y) = N_1u_1 + N_2u_2 + N_3u_3, \qquad \textbf{(9.6.9)}$$

where

$$N_1 = [(x_2y_3 - x_3y_2) + (y_2 - y_1)x + (x_3 - x_2)y]/(2A), \qquad \textbf{(9.6.10)}$$

$$N_2 = [(x_3y_1 - x_1y_3) + (y_3 - y_1)x + (x_1 - x_3)y]/(2A), \qquad \textbf{(9.6.11)}$$

and

$$N_3 = [(x_1y_2 - x_2y_1) + (y_1 - y_2)x + (x_2 - x_1)y]/(2A). \qquad \textbf{(9.6.12)}$$

Bilinear rectangular elements

Bilinear rectangular elements use the interpolation formula:

$$u(x, y) = N_1u_1 + N_2u_2 + N_3u_3 + N_4u_4. \qquad \textbf{(9.6.13)}$$

We wish Equation 9.6.13 to be linear in both x and y. Applying the same procedure as before, we have that

$$N_1 = (b - x)(a - y)/(4ab), \qquad N_2 = (b + x)(a - y)/(4ab), \qquad \textbf{(9.6.14)}$$

$$N_3 = (b + x)(a + y)/(4ab), \quad \text{and} \quad N_4 = (b - x)(a + y)/(4ab), \qquad \textbf{(9.6.15)}$$

where $2b$ and $2a$ are the length and height of the element, respectively.

Finite element formulation

Having introduced the two most common finite element representations, we are ready to apply them to Laplace's equation. As in the one dimension we develop the finite element formulation using the variational formulation (weak formulation) of the boundary-value problem by considering the integral

$$I = \iint_A w(x, y)\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) dx\,dy - \int_{\Gamma_e} w(x, y)\frac{\partial u}{\partial n}\,d\Gamma, \qquad \textbf{(9.6.16)}$$

where $w(x, y)$ is a weighting function and Γ_e is that portion of the boundary where the Dirichlet condition applies. Integrating Equation 9.6.16 by parts,

$$I = -\iint_A \left(\frac{\partial w}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial u}{\partial y}\right) dx\,dy + \int_{\Gamma_n} w(x, y)\frac{\partial u}{\partial n}\,d\Gamma, \qquad \textbf{(9.6.17)}$$

where Γ_n is that portion of the boundary where any Neumann condition exists.

For a linear triangular element, we have

$$[K^e]\mathbf{u}^e] = \iint_{\Omega^e} \left(\frac{\partial w}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial u}{\partial y} \right) dx\,dy \quad \textbf{(9.6.18)}$$

$$= \left\{ \iint_{\Omega^e} \sum_{n=1}^{3} \left[\left(\frac{\partial N_n}{\partial x} \right)^2 + \left(\frac{\partial N_n}{\partial y} \right)^2 \right] dx\,dy \right\} \mathbf{u}^e \quad \textbf{(9.6.19)}$$

where Ω^e is the element's domain. Upon substituting for N_1, N_2, and N_3 and carrying out the integrations,

$$[K^e] = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix}, \quad \textbf{(9.6.20)}$$

where

$$k_{11} = [(x_3 - x_2)^2 + (y_2 - y_3)^2]/(4A), \quad \textbf{(9.6.21)}$$

$$k_{12} = [(x_3 - x_2)(x_1 - x_3) + (y_2 - y_3)(y_3 - y_1)]/(4A), \quad \textbf{(9.6.22)}$$

$$k_{13} = [(x_3 - x_2)(x_2 - x_1) + (y_2 - y_3)(y_1 - y_2)]/(4A), \quad \textbf{(9.6.23)}$$

$$k_{21} = k_{12}, \quad \textbf{(9.6.24)}$$

$$k_{22} = [(x_1 - x_3)^2 + (y_3 - y_1)^2]/(4A), \quad \textbf{(9.6.25)}$$

$$k_{23} = [(x_1 - x_3)(x_2 - x_1) + (y_3 - y_1)(y_1 - y_2)]/(4A), \quad \textbf{(9.6.26)}$$

$$k_{31} = k_{13}, \quad \textbf{(9.6.27)}$$

$$k_{32} = k_{23}, \quad \textbf{(9.6.28)}$$

and

$$k_{33} = [(x_2 - x_1)^2 + (y_1 - y_2)^2]/(4A). \quad \textbf{(9.6.29)}$$

In the case of a bilinear element, we now have

$$[K^e]\mathbf{u}^e] = \left\{ \iint_{\Omega^e} \sum_{n=1}^{4} \left[\left(\frac{\partial N_n}{\partial x} \right)^2 + \left(\frac{\partial N_n}{\partial y} \right)^2 \right] dx\,dy \right\} \mathbf{u}^e. \quad \textbf{(9.6.30)}$$

Carrying out the integration, we obtain

$$K = \begin{pmatrix} k_{11} & k_{12} & k_{13} & k_{14} \\ k_{21} & k_{22} & k_{23} & k_{24} \\ k_{31} & k_{32} & k_{33} & k_{34} \\ k_{41} & k_{42} & k_{43} & k_{44} \end{pmatrix} \quad \textbf{(9.6.31)}$$

where

$$k_{11} = \frac{b^2 + a^2}{3ab}, \qquad k_{12} = \frac{b^2 - 2a^2}{6ab} \quad \textbf{(9.6.32)}$$

$$k_{13} = -\frac{b^2 + a^2}{6ab}, \qquad k_{14} = \frac{a^2 - 2b^2}{6ab} \quad \textbf{(9.6.33)}$$

$$k_{22} = k_{11}, \quad k_{23} = k_{14}, \quad k_{24} = k_{13}, \quad \textbf{(9.6.34)}$$

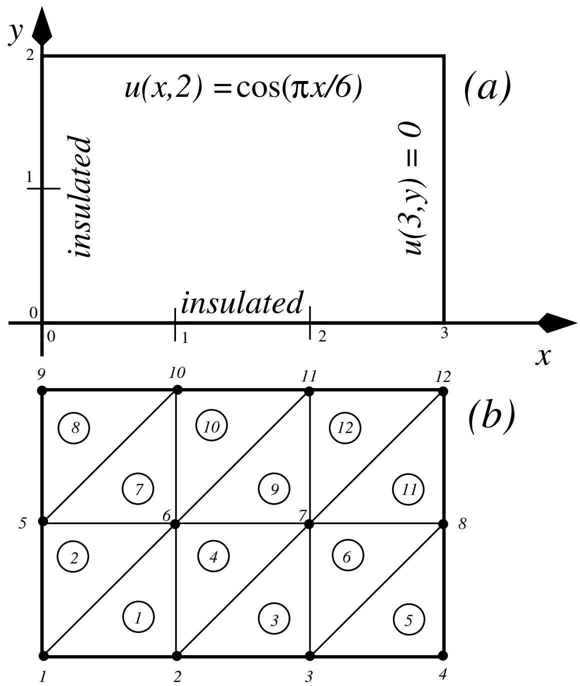


Figure 9.6.1: The (a) domain and (b) layout of nodes for Example 9.6.1. The numbers within the circles give the number for each triangular element.

and

$$k_{33} = k_{11}, \quad k_{34} = k_{12}, \quad k_{44} = k_{11}. \tag{9.6.35}$$

• Example 9.6.1

Let us illustrate the finite element method by finding the solution to Laplace's equation in a rectangular domain shown in Figure 9.6.1(a). The origin is taken at the lower left corner. The boundaries $x = 0$ and $y = 0$ are insulated, the boundary $x = 3$ is maintained at zero, and the boundary $y = 2$ is described by the equation $u(x, 2) = \cos(\pi x/6)$.

If we use 12 nodes, Figure 9.6.1(b) gives the global node number, element number, and element node numbers. The global node and element numberings are arbitrary. If we used some software package we would have to take care that we follow their convention. The element node numbering scheme used here is consistent with the element interpolation function, a counterclockwise system. According to the element node numbering scheme shown in Figure 9.6.1(b), all elements in the mesh fall into one of two geometric shapes: one with its base at the bottom of the element and another with its base at the top of the element. Our choice results in all of the elements having a common geometric shape, and thus the element coefficients must be computed only for a single element.

For a typical element of the mesh of triangles in Figure 9.6.1(b), the element coefficient matrix is

$$[K^e] = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}. \tag{9.6.36}$$

The element matrix is independent of the size of the element, as long as the element is a right-angled triangle where the base equals height.

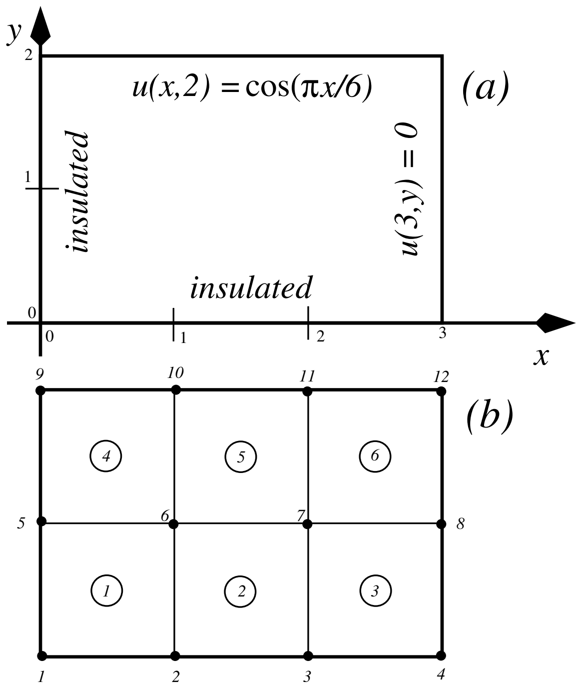


Figure 9.6.2: The (a) domain and (b) layout of nodes for Example 9.6.2. The numbers within the circles give the number for each rectangular element.

The boundary conditions require that $u_4 = u_8 = u_{12} = 0$, $u_9 = 1$, $u_{10} = \sqrt{3}/2$, $u_{11} = 1/2$, and there is no heat flow at the insulated boundary. Consequently the line integral vanishes and $[K^e]\mathbf{u}^e = \mathbf{0}$.

We first write the six finite element equations for the six unknown primary variables. These equations come from nodes 1, 2, 3, 5, 6, and 7:

$$\begin{pmatrix} 2 & -1 & 0 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 & -2 & 0 \\ 0 & -1 & 4 & 0 & 0 & -2 \\ -1 & 0 & 0 & 4 & -2 & 0 \\ 0 & -2 & 0 & -2 & 8 & -2 \\ 0 & 0 & -2 & 0 & -2 & 8 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_5 \\ u_6 \\ u_7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ \sqrt{3} \\ 1 \end{pmatrix}. \tag{9.6.37}$$

The solution to these equations is $u_1 = 0.6362$, $u_2 = 0.5510$, $u_3 = 0.3181$, $u_5 = 0.7214$, $u_6 = 0.6248$, and $u_7 = 0.3607$. The exact solution is $u_1^{exact} = 0.6249$, $u_2^{exact} = 0.5412$, $u_3^{exact} = 0.3124$, $u_5^{exact} = 0.7125$, $u_6^{exact} = 0.6171$, and $u_7^{exact} = 0.3563$. □

• Example 9.6.2

Let us redo Example 9.6.3 using rectangular elements. Figure 9.6.2 illustrates the grid now. The stiffness matrix for an element now becomes

$$[K^e] = \frac{1}{6} \begin{pmatrix} 4 & -1 & -2 & -1 \\ -1 & 4 & -1 & -2 \\ -2 & -1 & 4 & -1 \\ -1 & -2 & -1 & 4 \end{pmatrix}. \tag{9.6.38}$$

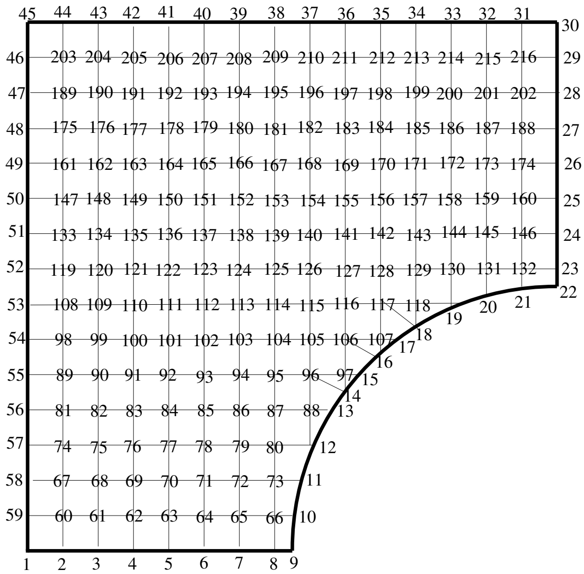


Figure 9.6.3: One of the possible nodal maps for the project.

If we consider only the number of nodes, the present mesh, shown in Figure 9.6.2(b), is the same as the triangular mesh that we used in Example 9.6.1. Hence the boundary conditions are unchanged here. The six finite elements equations for the unknown temperatures u_1, u_2, u_3, u_5, u_6, and u_7 are

$$\begin{pmatrix} 4 & -1 & 0 & -1 & -2 & 0 \\ -1 & 8 & -1 & -2 & -2 & -2 \\ 0 & -1 & 8 & 0 & -2 & -2 \\ -1 & -2 & 0 & 8 & -2 & 0 \\ -2 & -2 & -2 & -2 & 16 & -2 \\ 0 & -2 & -2 & 0 & -2 & 16 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_5 \\ u_6 \\ u_7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1+\sqrt{3} \\ 3+\sqrt{3} \\ 1+\sqrt{3} \end{pmatrix}. \qquad \textbf{(9.6.39)}$$

The solution to Equation 9.6.38 is $u_1 = 0.6128$, $u_2 = 0.5307$, $u_3 = 0.3064$, $u_5 = 0.7030$, $u_6 = 0.6088$, and $u_7 = 0.3515$. Note that our present results are *not* as accurate as those in the previous example. This occurs because there are only one-half of the elements as in the triangular element formulation.

Project: Solving Laplace's Equation Using Finite Elements

In this project you will use finite elements to solve Laplace's equation. The domain is square with a quarter circle removed from the lower right side of the square. Along the quarter circle, the solution equals one while along the straight edges the solution equals zero.

Although you could write your own MATLAB code, there already exist several codes[36] that are available online.[37] Consequently your principal task will be to construct the triangular and quadrilateral elements by choosing reasonable nodal positions. With those

[36] See, for example, Alberty, J., C. Carstensen, and S. A. Funken, 1999: Remarks about 50 lines of Matlab: Short finite element implementation. *Numer. Algorithms*, **20**, 117–137. This paper includes the MATLAB code that you will need.

[37] For example, `http://people.sc.fsu.edu/~burkardt/m_src/fem_50/fem_50.html`

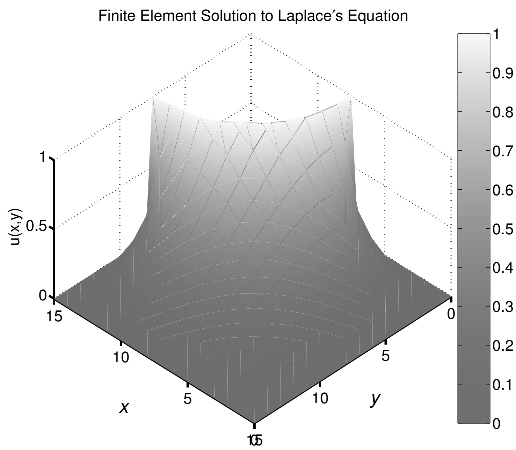


Figure 9.6.4: The finite element solution of Laplace's equation where the solution equals 0 along the straight edges and 1 along the quarter circle.

in place, the coding of the Dirichlet boundary conditions is straightforward. Figure 9.6.3 shows one possible nodal configuration. Figure 9.6.4 illustrates how your solution should appear.

Further Reading

Duffy, D. G., 2008: *Mixed Boundary Value Problems*. Chapman & Hall, 467 pp. This book gives a detailed account of mixed boundary value problems.

Koshlyakov, N. S., M. M. Smirnov, and E. B. Gliner, 1964: *Differential Equations of Mathematical Physics.* North-Holland Publishing, 701 pp. See Part II. Detailed presentation of mathematical techniques.

Morse, P. M., and H. Feshback, 1953: *Methods of Theoretical Physics.* McGraw-Hill Book Co., 997 pp. Chapter 10 is devoted to solving both Laplace's and Poisson's equations.

Chapter 10
Complex Variables

The theory of complex variables was originally developed by mathematicians as an aid in understanding functions. Functions of a complex variable enjoy many powerful properties that their real counterparts do not. That is *not* why we will study them. For us they provide the keys for the complete mastery of transform methods and differential equations.

In this chapter all of our work points to one objective: integration on the complex plane by the method of residues. For this reason we minimize discussions of limits and continuity, which play such an important role in conventional complex variables, in favor of the computational aspects. We begin by introducing some simple facts about complex variables. Then we progress to differential and integral calculus on the complex plane.

10.1 COMPLEX NUMBERS

A *complex number* is any number of the form $a+bi$, where a and b are real and $i = \sqrt{-1}$. We denote any member of a *set* of complex numbers by the *complex variable* $z = x + iy$. The real part of z, usually denoted by $\Re(z)$, is x while the imaginary part of z, $\Im(z)$, is y. The *complex conjugate*, $\overline{z}$ or z^*, of the complex number $a + bi$ is $a - bi$.

Complex numbers obey the fundamental rules of algebra. Thus, two complex numbers $a + bi$ and $c + di$ are equal if and only if $a = c$ and $b = d$. Just as real numbers have the fundamental operations of addition, subtraction, multiplication, and division, so too do complex numbers. These operations are defined:

Addition

$$(a + bi) + (c + di) = (a + c) + (b + d)i \tag{10.1.1}$$

Subtraction

$$(a + bi) - (c + di) = (a - c) + (b - d)i \tag{10.1.2}$$

Multiplication

$$(a+bi)(c+di) = ac + bci + adi + i^2bd = (ac-bd) + (ad+bc)i \tag{10.1.3}$$

Division

$$\frac{a+bi}{c+di} = \frac{a+bi}{c+di}\,\frac{c-di}{c-di} = \frac{ac-adi+bci-bdi^2}{c^2+d^2} = \frac{ac+bd+(bc-ad)i}{c^2+d^2}. \tag{10.1.4}$$

The *absolute value* or *modulus* of a complex number $a+bi$, written $|a+bi|$, equals $\sqrt{a^2+b^2}$. Additional properties include:

$$|z_1z_2z_3\cdots z_n| = |z_1||z_2||z_3|\cdots|z_n| \tag{10.1.5}$$

$$|z_1/z_2| = |z_1|/|z_2| \quad \text{if} \quad z_2 \neq 0 \tag{10.1.6}$$

$$|z_1+z_2+z_3+\cdots+z_n| \le |z_1|+|z_2|+|z_3|+\cdots+|z_n| \tag{10.1.7}$$

and

$$|z_1+z_2| \ge |z_1| - |z_2|. \tag{10.1.8}$$

The use of inequalities with complex variables has meaning only when they involve absolute values.

It is often useful to plot the complex number $x+iy$ as a point (x,y) in the xy-plane, now called the *complex plane*. Figure 10.1.1 illustrates this representation.

This geometrical interpretation of a complex number suggests an alternative method of expressing a complex number: the polar form. From the polar representation of x and y,

$$x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta), \tag{10.1.9}$$

where $r=\sqrt{x^2+y^2}$ is the *modulus*, *amplitude*, or *absolute value* of z and θ is the *argument* or *phase*, we have that

$$z = x+iy = r[\cos(\theta)+i\sin(\theta)]. \tag{10.1.10}$$

However, from the Taylor expansion of the exponential in the real case,

$$e^{i\theta} = \sum_{k=0}^{\infty}\frac{(\theta i)^k}{k!}. \tag{10.1.11}$$

Expanding Equation 10.1.11,

$$e^{i\theta} = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots\right) \tag{10.1.12}$$

$$= \cos(\theta) + i\sin(\theta). \tag{10.1.13}$$

Equation 10.1.13 is *Euler's formula*. Consequently, we may express Equation 10.1.10 as

$$z = re^{i\theta}, \tag{10.1.14}$$

which is the *polar form* of a complex number. Furthermore, because

$$z^n = r^ne^{in\theta} \tag{10.1.15}$$

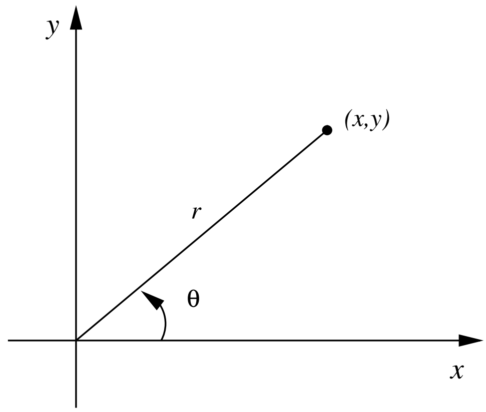


Figure 10.1.1: The complex plane.

by the law of exponents,

$$z^n = r^n[\cos(n\theta) + i\sin(n\theta)]. \tag{10.1.16}$$

Equation 10.1.16 is *De Moivre's theorem.*

• Example 10.1.1

Let us simplify the following complex number:

$$\frac{3-2i}{-1+i} = \frac{3-2i}{-1+i} \times \frac{-1-i}{-1-i} = \frac{-3-3i+2i+2i^2}{1+1} = \frac{-5-i}{2} = -\frac{5}{2} - \frac{i}{2}. \tag{10.1.17}$$

□

• Example 10.1.2

Let us reexpress the complex number $-\sqrt{6} - i\sqrt{2}$ in polar form. From Equation 10.1.9 $r = \sqrt{6+2}$ and $\theta = \tan^{-1}(b/a) = \tan^{-1}(1/\sqrt{3}) = \pi/6$ or $7\pi/6$. Because $-\sqrt{6} - i\sqrt{2}$ lies in the third quadrant of the complex plane, $\theta = 7\pi/6$ and

$$-\sqrt{6} - i\sqrt{2} = 2\sqrt{2}e^{7\pi i/6}. \tag{10.1.18}$$

Note that Equation 10.1.18 is not a unique representation because $\pm 2n\pi$ may be added to $7\pi/6$ and we still have the same complex number since

$$e^{i(\theta \pm 2n\pi)} = \cos(\theta \pm 2n\pi) + i\sin(\theta \pm 2n\pi) = \cos(\theta) + i\sin(\theta) = e^{i\theta}. \tag{10.1.19}$$

For uniqueness we often choose $n = 0$ and define this choice as the *principal branch.* Other branches correspond to different values of n. □

• Example 10.1.3

Find the curve described by the equation $|z - z_0| = a$.

From the definition of the absolute value,

$$\sqrt{(x-x_0)^2 + (y-y_0)^2} = a \tag{10.1.20}$$

or

$$(x - x_0)^2 + (y - y_0)^2 = a^2. \tag{10.1.21}$$

Equation 10.1.21, and hence $|z - z_0| = a$, describes a circle of radius a with its center located at (x_0, y_0). Later on, we shall use equations such as this to describe curves in the complex plane. □

• **Example 10.1.4**

As an example in manipulating complex numbers, let us show that

$$\left|\frac{a+bi}{b+ai}\right| = 1. \tag{10.1.22}$$

We begin by simplifying

$$\frac{a+bi}{b+ai} = \frac{a+bi}{b+ai} \times \frac{b-ai}{b-ai} = \frac{2ab}{a^2+b^2} + \frac{b^2-a^2}{a^2+b^2}i. \tag{10.1.23}$$

Therefore,

$$\left|\frac{a+bi}{b+ai}\right| = \sqrt{\frac{4a^2b^2}{(a^2+b^2)^2} + \frac{b^4-2a^2b^2+a^4}{(a^2+b^2)^2}} = \sqrt{\frac{a^4+2a^2b^2+b^4}{(a^2+b^2)^2}} = 1. \tag{10.1.24}$$

MATLAB can also be used to solve this problem. Typing the commands

```
>> syms a b real
>> abs((a+b*i)/(b+a*i))
```

yields

```
ans =
1
```

Note that you must declare a and b real in order to get the final result.

Problems

Simplify the following complex numbers. Represent the solution in the Cartesian form $a+bi$. Check your answers using MATLAB.

1. $\dfrac{5i}{2+i}$

2. $\dfrac{5+5i}{3-4i} + \dfrac{20}{4+3i}$

3. $\dfrac{1+2i}{3-4i} + \dfrac{2-i}{5i}$

4. $(1-i)^4$

5. $i(1-i\sqrt{3})(\sqrt{3}+i)$

Represent the following complex numbers in polar form:

6. $-i$

7. -4

8. $2+2\sqrt{3}\,i$

9. $-5+5i$

10. $2-2i$

11. $-1+\sqrt{3}\,i$

12. By the law of exponents, $e^{i(\alpha+\beta)} = e^{i\alpha}e^{i\beta}$. Use Euler's formula to obtain expressions for $\cos(\alpha+\beta)$ and $\sin(\alpha+\beta)$ in terms of sines and cosines of α and β.

13. Using the property that $\sum_{n=0}^{N} q^n = (1-q^{N+1})/(1-q)$ and the geometric series $\sum_{n=0}^{N} e^{int}$, obtain the following sums of trigonometric functions:

$$\sum_{n=0}^{N} \cos(nt) = \cos\left(\frac{Nt}{2}\right)\frac{\sin[(N+1)t/2]}{\sin(t/2)} \quad \text{and} \quad \sum_{n=1}^{N} \sin(nt) = \sin\left(\frac{Nt}{2}\right)\frac{\sin[(N+1)t/2]}{\sin(t/2)}.$$

These results are often called *Lagrange's trigonometric identities.*

14. (a) Using the property that $\sum_{n=0}^{\infty} q^n = 1/(1-q)$, if $|q| < 1$, and the geometric series $\sum_{n=0}^{\infty} \epsilon^n e^{int}$, $|\epsilon| < 1$, show that

$$\sum_{n=0}^{\infty} \epsilon^n \cos(nt) = \frac{1-\epsilon\cos(t)}{1+\epsilon^2-2\epsilon\cos(t)} \quad \text{and} \quad \sum_{n=1}^{\infty} \epsilon^n \sin(nt) = \frac{\epsilon\sin(t)}{1+\epsilon^2-2\epsilon\cos(t)}.$$

(b) Let $\epsilon = e^{-a}$, where $a > 0$. Show that

$$2\sum_{n=1}^{\infty} e^{-na}\sin(nt) = \frac{\sin(t)}{\cosh(a)-\cos(t)}.$$

10.2 FINDING ROOTS

The concept of finding roots of a number, which is rather straightforward in the case of real numbers, becomes more difficult in the case of complex numbers. By finding the *roots* of a complex number, we wish to find all of the solutions w of the equation $w^n = z$, where n is a positive integer for a given z.

We begin by writing z in the polar form:

$$z = re^{i\varphi}, \tag{10.2.1}$$

while we write

$$w = Re^{i\Phi} \tag{10.2.2}$$

for the unknown. Consequently,

$$w^n = R^n e^{in\Phi} = re^{i\varphi} = z. \tag{10.2.3}$$

We satisfy Equation 10.2.3 if

$$R^n = r \quad \text{and} \quad n\Phi = \varphi + 2k\pi, \quad k = 0, \pm1, \pm2, \ldots, \tag{10.2.4}$$

because the addition of any multiple of 2π to the argument is also a solution. Thus, $R = r^{1/n}$, where R is the uniquely determined real positive root, and

$$\Phi_k = \frac{\varphi}{n} + \frac{2\pi k}{n}, \qquad k = 0, \pm1, \pm2, \ldots. \tag{10.2.5}$$

Because $w_k = w_{k\pm n}$, it is sufficient to take $k = 0, 1, 2, \ldots, n-1$. Therefore, there are exactly n solutions:

$$w_k = Re^{\Phi_k i} = r^{1/n}\exp\left[i\left(\frac{\varphi}{n} + \frac{2\pi k}{n}\right)\right] \tag{10.2.6}$$

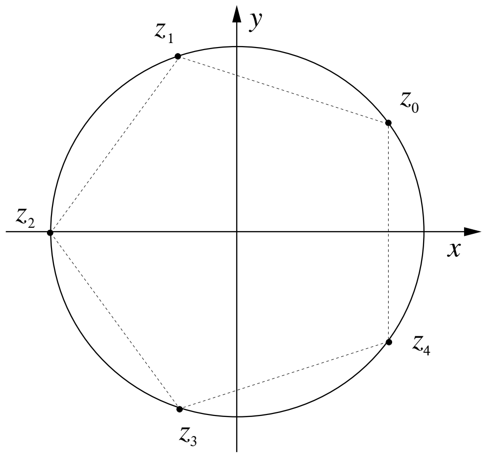


Figure 10.2.1: The zeros of $z^5 = -32$.

with $k = 0, 1, 2, \ldots, n-1$. They are the n roots of z. Geometrically we can locate these solutions w_k on a circle, centered at the point $(0, 0)$, with radius R and separated from each other by $2\pi/n$ radians. These roots also form the vertices of a regular polygon of n sides inscribed inside of a circle of radius R. (See Example 10.2.1.)

In summary, the method for finding the n roots of a complex number z_0 is as follows. First, write z_0 in its polar form: $z_0 = re^{i\varphi}$. Then multiply the polar form by $e^{2i\pi k}$. Using the law of exponents, take the $1/n$ power of both sides of the equation. Finally, using Euler's formula, evaluate the roots for $k = 0, 1, \ldots, n-1$.

• Example 10.2.1

Let us find all of the values of z for which $z^5 = -32$ and locate these values on the complex plane.

Because

$$-32 = 32e^{\pi i} = 2^5 e^{\pi i}, \tag{10.2.7}$$

$$z_k = 2\exp\left(\frac{\pi i}{5} + \frac{2\pi ik}{5}\right), \qquad k = 0, 1, 2, 3, 4, \tag{10.2.8}$$

or

$$z_0 = 2\exp\left(\frac{\pi i}{5}\right) = 2\left[\cos\left(\frac{\pi}{5}\right) + i\sin\left(\frac{\pi}{5}\right)\right], \tag{10.2.9}$$

$$z_1 = 2\exp\left(\frac{3\pi i}{5}\right) = 2\left[\cos\left(\frac{3\pi}{5}\right) + i\sin\left(\frac{3\pi}{5}\right)\right], \tag{10.2.10}$$

$$z_2 = 2e^{\pi i} = -2, \tag{10.2.11}$$

$$z_3 = 2\exp\left(\frac{7\pi i}{5}\right) = 2\left[\cos\left(\frac{7\pi}{5}\right) + i\sin\left(\frac{7\pi}{5}\right)\right] \tag{10.2.12}$$

and

$$z_4 = 2\exp\left(\frac{9\pi i}{5}\right) = 2\left[\cos\left(\frac{9\pi}{5}\right) + i\sin\left(\frac{9\pi}{5}\right)\right]. \tag{10.2.13}$$

Figure 10.2.1 shows the location of these roots in the complex plane. □

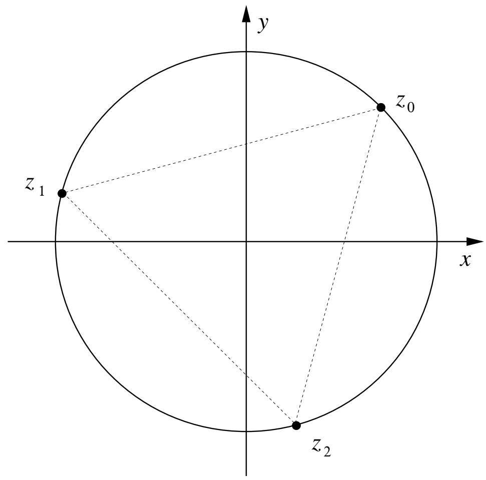


Figure 10.2.2: The zeros of $z^3 = -1 + i$.

• Example 10.2.2

Let us find the cube roots of $-1 + i$ and locate them graphically.
Because $-1 + i = \sqrt{2}\,\exp(3\pi i/4)$,

$$z_k = 2^{1/6}\exp\left(\frac{\pi i}{4} + \frac{2i\pi k}{3}\right), \qquad k = 0, 1, 2, \tag{10.2.14}$$

or

$$z_0 = 2^{1/6}\exp\left(\frac{\pi i}{4}\right) = 2^{1/6}\left[\cos\left(\frac{\pi}{4}\right) + i\sin\left(\frac{\pi}{4}\right)\right], \tag{10.2.15}$$

$$z_1 = 2^{1/6}\exp\left(\frac{11\pi i}{12}\right) = 2^{1/6}\left[\cos\left(\frac{11\pi}{12}\right) + i\sin\left(\frac{11\pi}{12}\right)\right], \tag{10.2.16}$$

and

$$z_2 = 2^{1/6}\exp\left(\frac{19\pi i}{12}\right) = 2^{1/6}\left[\cos\left(\frac{19\pi}{12}\right) + i\sin\left(\frac{19\pi}{12}\right)\right]. \tag{10.2.17}$$

Figure 10.2.2 gives the location of these zeros on the complex plane. □

• Example 10.2.3

The routine `solve` in MATLAB can also be used to compute the roots of complex numbers. For example, let us find all of the roots of $z^4 = -a^4$.

The MATLAB commands are as follows:

```
>> syms a z
>> solve(z^4+a^4)
```

This yields the solution

```
ans=
[  (1/2*2^(1/2)+1/2*i*2^(1/2))*a]
[ (-1/2*2^(1/2)+1/2*i*2^(1/2))*a]
[  (1/2*2^(1/2)-1/2*i*2^(1/2))*a]
[ (-1/2*2^(1/2)-1/2*i*2^(1/2))*a]
```

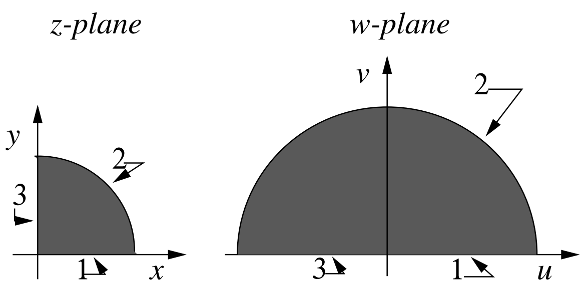


Figure 10.3.1: The complex function $w = z^2$.

Problems

Extract all of the possible roots of the following complex numbers. Verify your answer using MATLAB.

1. $8^{1/6}$ 2. $(-1)^{1/3}$ 3. $(-i)^{1/3}$ 4. $(-27i)^{1/6}$

5. Find algebraic expressions for the square roots of $a - bi$, where $a > 0$ and $b > 0$.

6. Find all of the roots for the algebraic equation $z^4 - 3iz^2 - 2 = 0$. Then check your answer using `solve` in MATLAB.

7. Find all of the roots for the algebraic equation $z^4 + 6iz^2 + 16 = 0$. Then check your answer using `solve` in MATLAB.

10.3 THE DERIVATIVE IN THE COMPLEX PLANE: THE CAUCHY-RIEMANN EQUATIONS

In the previous two sections, we introduced complex arithmetic. We are now ready for the concept of function as it applies to complex variables.

We already defined the complex variable $z = x+iy$, where x and y are variable. We now introduce another complex variable $w = u+iv$ so that for each value of z there corresponds a value of $w = f(z)$. From all of the possible complex functions that we might invent, we focus on those functions where for each z there is one, and only one, value of w. These functions are *single-valued*. They differ from functions such as the square root, logarithm, and inverse sine and cosine, where there are multiple answers for each z. These *multivalued functions* do arise in various problems. However, they are beyond the scope of this book and we shall always assume that we are dealing with single-valued functions.

A popular method for representing a complex function involves drawing some closed domain in the z-plane and then showing the corresponding domain in the w-plane. This procedure is called *mapping* and the z-plane illustrates the *domain* of the function while the w-plane illustrates its *image* or *range*. Figure 10.3.1 shows the z-plane and w-plane for $w = z^2$; a pie-shaped wedge in the z-plane maps into a semicircle on the w-plane.

• Example 10.3.1

Given the complex function $w = e^{-z^2}$, let us find the corresponding $u(x,y)$ and $v(x,y)$. From Euler's formula,

$$w = e^{-z^2} = e^{-(x+iy)^2} = e^{y^2-x^2}e^{-2ixy} = e^{y^2-x^2}[\cos(2xy) - i\sin(2xy)]. \qquad \textbf{(10.3.1)}$$

Therefore, by inspection,

$$u(x, y) = e^{y^2 - x^2} \cos(2xy), \qquad \text{and} \qquad v(x, y) = -e^{y^2 - x^2} \sin(2xy). \tag{10.3.2}$$

Note that there is no i in the expression for $v(x, y)$. The function $w = f(z)$ is single-valued because for each distinct value of z, there is a unique value of $u(x, y)$ and $v(x, y)$. □

• **Example 10.3.2**

As counterpoint, let us show that $w = \sqrt{z}$ is a multivalued function.

We begin by writing $z = re^{i\theta + 2\pi ik}$, where $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$. Then,

$$w_k = \sqrt{r} e^{i\theta/2 + \pi ik}, \qquad k = 0, 1, \tag{10.3.3}$$

or

$$w_0 = \sqrt{r}\,[\cos(\theta/2) + i \sin(\theta/2)] \qquad \text{and} \qquad w_1 = -w_0. \tag{10.3.4}$$

Therefore,

$$u_0(x, y) = \sqrt{r} \cos(\theta/2), \qquad v_0(x, y) = \sqrt{r} \sin(\theta/2), \tag{10.3.5}$$

and

$$u_1(x, y) = -\sqrt{r} \cos(\theta/2), \qquad v_1(x, y) = -\sqrt{r} \sin(\theta/2). \tag{10.3.6}$$

Each solution w_0 or w_1 is a *branch* of the multivalued function $\sqrt{z}$. We can make $\sqrt{z}$ single-valued by restricting ourselves to a single branch, say w_0. In that case, the $\Re(w) > 0$ if we restrict $-\pi < \theta < \pi$. Although this is not the only choice that we could have made, it is a popular one. For example, most digital computers use this definition in their complex square root function. The point here is our ability to make a multivalued function single-valued by defining a particular branch. □

Although the requirement that a complex function be single-valued is important, it is still too general and would cover all functions of two real variables. To have a useful theory, we must introduce additional constraints. Because an important property associated with most functions is the ability to take their derivative, let us examine the derivative in the complex plane.

Following the definition of a derivative for a single real variable, the derivative of a complex function $w = f(z)$ is defined as

$$\frac{dw}{dz} = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}. \tag{10.3.7}$$

A function of a complex variable that has a derivative at every point within a region of the complex plane is said to be *analytic* (or *regular* or *holomorphic*) over that region. If the function is analytic everywhere in the complex plane, it is *entire*.

Because the derivative is defined as a limit and limits are well behaved with respect to elementary algebraic operations, the following operations carry over from elementary calculus:

$$\frac{d}{dz}\left[cf(z)\right] = cf'(z), \qquad c \text{ a constant} \tag{10.3.8}$$

$$\frac{d}{dz}\left[f(z) \pm g(z)\right] = f'(z) \pm g'(z) \tag{10.3.9}$$

$$\frac{d}{dz}\left[f(z)g(z)\right] = f'(z)g(z) + f(z)g'(z) \tag{10.3.10}$$

$$\frac{d}{dz}\left[\frac{f(z)}{g(z)}\right] = \frac{g(z)f'(z) - g'(z)f(z)}{g^2(z)} \tag{10.3.11}$$

$$\frac{d}{dz}\left\{f[g(z)]\right\} = f'[g(z)]g'(z), \qquad \text{the chain rule.} \tag{10.3.12}$$

Another important property that carries over from real variables is l'Hôspital's rule: Let $f(z)$ and $g(z)$ be analytic at z_0, where $f(z)$ has a zero[1] of order m and $g(z)$ has a zero of order n. Then, if $m > n$,

$$\lim_{z\to z_0} \frac{f(z)}{g(z)} = 0; \tag{10.3.13}$$

if $m = n$,

$$\lim_{z\to z_0} \frac{f(z)}{g(z)} = \frac{f^{(m)}(z_0)}{g^{(m)}(z_0)}; \tag{10.3.14}$$

and if $m < n$,

$$\lim_{z\to z_0} \frac{f(z)}{g(z)} = \infty. \tag{10.3.15}$$

• Example 10.3.3

Let us evaluate $\lim_{z\to i}(z^{10}+1)/(z^6+1)$. From l'Hôspital's rule,

$$\lim_{z\to i} \frac{z^{10}+1}{z^6+1} = \lim_{z\to i} \frac{10z^9}{6z^5} = \frac{5}{3}\lim_{z\to i} z^4 = \frac{5}{3}. \tag{10.3.16}$$

□

So far, we introduced the derivative and some of its properties. But how do we actually know whether a function is analytic or how do we compute its derivative? At this point we must develop some relationships involving the known quantities $u(x,y)$ and $v(x,y)$.

We begin by returning to the definition of the derivative. Because $\Delta z = \Delta x + i\Delta y$, there is an infinite number of different ways of approaching the limit $\Delta z \to 0$. Uniqueness of that limit requires that Equation 10.3.7 must be independent of the manner in which Δz approaches zero. A simple example is to take Δz in the x-direction so that $\Delta z = \Delta x$; another is to take Δz in the y-direction so that $\Delta z = i\Delta y$. These examples yield

$$\frac{dw}{dz} = \lim_{\Delta z\to 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta x\to 0} \frac{\Delta u + i\Delta v}{\Delta x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} \tag{10.3.17}$$

and

$$\frac{dw}{dz} = \lim_{\Delta z\to 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta y\to 0} \frac{\Delta u + i\Delta v}{i\Delta y} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}. \tag{10.3.18}$$

[1] An analytic function $f(z)$ has a zero of order m at z_0 if and only if $f(z_0) = f'(z_0) = \cdots = f^{(m-1)}(z_0) = 0$ and $f^{(m)}(z_0) \neq 0$.

Although educated as an engineer, Augustin-Louis Cauchy (1789–1857) would become a mathematician's mathematician, publishing 789 papers and 7 books in the fields of pure and applied mathematics. His greatest writings established the discipline of mathematical analysis as he refined the notions of limit, continuity, function, and convergence. It was this work on analysis that led him to develop complex function theory via the concept of residues. (Portrait courtesy of the Archives de l'Académie des sciences, Paris.)

In both cases we are approaching zero from the positive side. For the limit to be unique and independent of path, Equation 10.3.17 must equal Equation 10.3.18, or

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{10.3.19}$$

These equations that u and v must both satisfy are the *Cauchy-Riemann* equations. They are necessary but not sufficient to ensure that a function is differentiable. The following example illustrates this.

• Example 10.3.4

Consider the complex function

$$w = \begin{cases} z^5/|z|^4, & z \neq 0 \\ 0, & z = 0. \end{cases} \tag{10.3.20}$$

The derivative at $z = 0$ is given by

$$\frac{dw}{dz} = \lim_{\Delta z \to 0} \frac{(\Delta z)^5/|\Delta z|^4 - 0}{\Delta z} = \lim_{\Delta z \to 0} \frac{(\Delta z)^4}{|\Delta z|^4}, \tag{10.3.21}$$

Despite his short life, (Georg Friedrich) Bernhard Riemann's (1826–1866) mathematical work contained many imaginative and profound concepts. It was in his doctoral thesis on complex function theory (1851) that he introduced the Cauchy-Riemann differential equations. Riemann's later work dealt with the definition of the integral and the foundations of geometry and non-Euclidean (elliptic) geometry. (Portrait courtesy of Photo AKG, London, with permission.)

provided that this limit exists. However, this limit does not exist because, in general, the numerator depends upon the path used to approach zero. For example, if $\Delta z = re^{\pi i/4}$ with $r \to 0$, $dw/dz = -1$. On the other hand, if $\Delta z = re^{\pi i/2}$ with $r \to 0$, $dw/dz = 1$.

Are the Cauchy-Riemann equations satisfied in this case? To check this, we first compute

$$u_x(0,0) = \lim_{\Delta x \to 0} \left(\frac{\Delta x}{|\Delta x|} \right)^4 = 1, \qquad v_y(0,0) = \lim_{\Delta y \to 0} \left(\frac{i\Delta y}{|\Delta y|} \right)^4 = 1, \tag{10.3.22}$$

$$u_y(0,0) = \lim_{\Delta y \to 0} \Re \left[\frac{(i\Delta y)^5}{\Delta y|\Delta y|^4} \right] = 0, \quad \text{and} \quad v_x(0,0) = \lim_{\Delta x \to 0} \Im \left[\left(\frac{\Delta x}{|\Delta x|} \right)^4 \right] = 0. \tag{10.3.23}$$

Hence, the Cauchy-Riemann equations are satisfied at the origin. Thus, even though the derivative is not uniquely defined, Equation 10.3.21 happens to have the same value for paths taken along the coordinate axes so that the Cauchy-Riemann equations are satisfied. □

In summary, if a function is differentiable at a point, the Cauchy-Riemann equations hold. Similarly, if the Cauchy-Riemann equations are not satisfied at a point, then the function is not differentiable at that point. This is one of the important uses of the Cauchy-Riemann equations: the location of nonanalytic points. Isolated nonanalytic points of an otherwise analytic function are called *isolated singularities*. Functions that contain isolated singularities are called *meromorphic*.

The Cauchy-Riemann condition can be modified so that it is sufficient for the derivative to exist. Let us require that u_x, u_y, v_x, and v_y be continuous in some region surrounding a point z_0 and satisfy the Cauchy-Riemann equations there. Then

$$
\begin{aligned}
f(z) - f(z_0) &= [u(z) - u(z_0)] + i[v(z) - v(z_0)] && \textbf{(10.3.24)}\\
&= [u_x(z_0)(x - x_0) + u_y(z_0)(y - y_0) + \epsilon_1(x - x_0) + \epsilon_2(y - y_0)] \\
&\quad + i[v_x(z_0)(x - x_0) + v_y(z_0)(y - y_0) + \epsilon_3(x - x_0) + \epsilon_4(y - y_0)] && \textbf{(10.3.25)}\\
&= [u_x(z_0) + iv_x(z_0)](z - z_0) + (\epsilon_1 + i\epsilon_3)(x - x_0) \\
&\quad + (\epsilon_2 + i\epsilon_4)(y - y_0), && \textbf{(10.3.26)}
\end{aligned}
$$

where we used the Cauchy-Riemann equations and $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 \to 0$ as $\Delta x, \Delta y \to 0$. Hence,

$$
f'(z_0) = \lim_{\Delta z \to 0} \frac{f(z) - f(z_0)}{\Delta z} = u_x(z_0) + iv_x(z_0), \qquad \textbf{(10.3.27)}
$$

because $|\Delta x| \leq |\Delta z|$ and $|\Delta y| \leq |\Delta z|$. Using Equation 10.3.27 and the Cauchy-Riemann equations, we can obtain the derivative from any of the following formulas:

$$
\boxed{\frac{dw}{dz} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y},} \qquad \textbf{(10.3.28)}
$$

and

$$
\boxed{\frac{dw}{dz} = \frac{\partial v}{\partial y} + i\frac{\partial v}{\partial x} = \frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y}.} \qquad \textbf{(10.3.29)}
$$

Furthermore, $f'(z_0)$ is continuous because the partial derivatives are.

• Example 10.3.5

Let us show that $\sin(z)$ is an entire function.

$$
\begin{aligned}
w &= \sin(z) && \textbf{(10.3.30)}\\
u + iv &= \sin(x + iy) = \sin(x)\cos(iy) + \cos(x)\sin(iy) && \textbf{(10.3.31)}\\
&= \sin(x)\cosh(y) + i\cos(x)\sinh(y), && \textbf{(10.3.32)}
\end{aligned}
$$

because

$$
\cos(iy) = \tfrac{1}{2}\left[e^{i(iy)} + e^{-i(iy)}\right] = \tfrac{1}{2}\left[e^{y} + e^{-y}\right] = \cosh(y), \qquad \textbf{(10.3.33)}
$$

and

$$
\sin(iy) = \tfrac{1}{2i}\left[e^{i(iy)} - e^{-i(iy)}\right] = -\tfrac{1}{2i}\left[e^{y} - e^{-y}\right] = i\sinh(y), \qquad \textbf{(10.3.34)}
$$

so that

$$
u(x, y) = \sin(x)\cosh(y), \quad \text{and} \quad v(x, y) = \cos(x)\sinh(y). \qquad \textbf{(10.3.35)}
$$

Differentiating both $u(x, y)$ and $v(x, y)$ with respect to x and y, we have that

$$
\frac{\partial u}{\partial x} = \cos(x)\cosh(y), \qquad \frac{\partial u}{\partial y} = \sin(x)\sinh(y), \qquad \textbf{(10.3.36)}
$$

$$\frac{\partial v}{\partial x} = -\sin(x)\sinh(y), \qquad \frac{\partial v}{\partial y} = \cos(x)\cosh(y), \tag{10.3.37}$$

and $u(x,y)$ and $v(x,y)$ satisfy the Cauchy-Riemann equations for all values of x and y. Furthermore, u_x, u_y, v_x, and v_y are continuous for all x and y. Therefore, the function $w = \sin(z)$ is an entire function. □

• Example 10.3.6

Consider the function $w = 1/z$. Then

$$w = u + iv = \frac{1}{x+iy} = \frac{x}{x^2+y^2} - \frac{iy}{x^2+y^2}. \tag{10.3.38}$$

Therefore,

$$u(x,y) = \frac{x}{x^2+y^2}, \qquad \text{and} \qquad v(x,y) = -\frac{y}{x^2+y^2}. \tag{10.3.39}$$

Now

$$\frac{\partial u}{\partial x} = \frac{(x^2+y^2)-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}, \tag{10.3.40}$$

$$\frac{\partial v}{\partial y} = -\frac{(x^2+y^2)-2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{\partial u}{\partial x}, \tag{10.3.41}$$

$$\frac{\partial v}{\partial x} = -\frac{0-2xy}{(x^2+y^2)^2} = \frac{2xy}{(x^2+y^2)^2}, \tag{10.3.42}$$

and

$$\frac{\partial u}{\partial y} = \frac{0-2xy}{(x^2+y^2)^2} = -\frac{2xy}{(x^2+y^2)^2} = -\frac{\partial v}{\partial x}. \tag{10.3.43}$$

The function is analytic at all points except the origin because the function itself ceases to exist when both x and y are zero and the modulus of w becomes infinite. □

• Example 10.3.7

Let us find the derivative of $\sin(z)$.
Using Equation 10.3.28 and Equation 10.3.32,

$$\frac{d}{dz}\left[\sin(z)\right] = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \cos(x)\cosh(y) - i\sin(x)\sinh(y) = \cos(x+iy) = \cos(z). \tag{10.3.44}$$

Similarly,

$$\frac{d}{dz}\left(\frac{1}{z}\right) = \frac{y^2-x^2}{(x^2+y^2)^2} + \frac{2ixy}{(x^2+y^2)^2} = -\frac{1}{(x+iy)^2} = -\frac{1}{z^2}. \tag{10.3.45}$$

□

The results in the above examples are identical to those for z real. As we showed earlier, the fundamental rules of elementary calculus apply to complex differentiation. Consequently, it is usually simpler to apply those rules to find the derivative rather than breaking $f(z)$ down into its real and imaginary parts, applying either Equation 10.3.28 or Equation 10.3.29, and then putting everything back together.

An additional property of analytic functions follows by cross differentiating the Cauchy-Riemann equations, or

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y} = -\frac{\partial^2 u}{\partial y^2}, \quad \text{or} \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{10.3.46}$$

and

$$\frac{\partial^2 v}{\partial x^2} = -\frac{\partial^2 u}{\partial x \partial y} = -\frac{\partial^2 v}{\partial y^2}, \quad \text{or} \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0. \tag{10.3.47}$$

Any function that has continuous partial derivatives of second order and satisfies Laplace's equation, Equation 10.3.46 or Equation 10.3.47, is called a *harmonic function*. Because both $u(x, y)$ and $v(x, y)$ satisfy Laplace's equation if $f(z) = u + iv$ is analytic, $u(x, y)$ and $v(x, y)$ are called *conjugate harmonic functions*.

• Example 10.3.8

Given that $u(x, y) = e^{-x}[x \sin(y) - y \cos(y)]$, let us show that u is harmonic and find a conjugate harmonic function $v(x, y)$ such that $f(z) = u + iv$ is analytic.

Because

$$\frac{\partial^2 u}{\partial x^2} = -2e^{-x} \sin(y) + xe^{-x} \sin(y) - ye^{-x} \cos(y), \tag{10.3.48}$$

and

$$\frac{\partial^2 u}{\partial y^2} = -xe^{-x} \sin(y) + 2e^{-x} \sin(y) + ye^{-x} \cos(y), \tag{10.3.49}$$

it follows that $u_{xx} + u_{yy} = 0$. Therefore, $u(x, y)$ is harmonic. From the Cauchy-Riemann equations,

$$\frac{\partial v}{\partial y} = \frac{\partial u}{\partial x} = e^{-x} \sin(y) - xe^{-x} \sin(y) + ye^{-x} \cos(y), \tag{10.3.50}$$

and

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} = e^{-x} \cos(y) - xe^{-x} \cos(y) - ye^{-x} \sin(y). \tag{10.3.51}$$

Integrating Equation 10.3.50 with respect to y,

$$v(x, y) = ye^{-x} \sin(y) + xe^{-x} \cos(y) + g(x). \tag{10.3.52}$$

Using Equation 10.3.51,

$$\begin{aligned} v_x &= -ye^{-x} \sin(y) - xe^{-x} \cos(y) + e^{-x} \cos(y) + g'(x) \\ &= e^{-x} \cos(y) - xe^{-x} \cos(y) - ye^{-x} \sin(x). \end{aligned} \quad \textbf{(10.3.53}$$

Therefore, $g'(x) = 0$ or $g(x) = \text{constant}$. Consequently,

$$v(x, y) = e^{-x}[y \sin(y) + x \cos(y)] + \text{constant}. \tag{10.3.54}$$

Hence, for our real harmonic function $u(x, y)$, there are infinitely many harmonic conjugates $v(x, y)$, which differ from each other by an additive constant.

Problems

Show that the following functions are entire:

1. $f(z) = iz + 2$ 2. $f(z) = e^{-z}$ 3. $f(z) = z^3$ 4. $f(z) = \cosh(z)$

Find the derivative of the following functions:

5. $f(z) = (1 + z^2)^{3/2}$ 6. $f(z) = (z + 2z^{1/2})^{1/3}$ 7. $f(z) = (1 + 4i)z^2 - 3z - 2$

8. $f(z) = (2z - i)/(z + 2i)$ 9. $f(z) = (iz - 1)^{-3}$

Evaluate the following limits:

10. $\lim_{z \to i} \dfrac{z^2 - 2iz - 1}{z^4 + 2z^2 + 1}$ 11. $\lim_{z \to 0} \dfrac{z - \sin(z)}{z^3}$

12. Show that the function $f(z) = z^*$ is nowhere differentiable.

For each of the following $u(x, y)$, show that it is harmonic and then find a corresponding $v(x, y)$ such that $f(z) = u + iv$ is analytic.

13. $u(x, y) = x^2 - y^2$ 14. $u(x, y) = x^4 - 6x^2y^2 + y^4 + x$

15. $u(x, y) = x\cos(x)e^{-y} - y\sin(x)e^{-y}$ 16. $u(x, y) = (x^2 - y^2)\cos(y)e^x - 2xy\sin(y)e^x$

10.4 LINE INTEGRALS

So far, we discussed complex numbers, complex functions, and complex differentiation. We are now ready for integration.

Just as we have integrals involving real variables, we can define an integral that involves complex variables. Because the z-plane is two-dimensional, there is clearly greater freedom in what we mean by a complex integral. For example, we might ask whether the integral of some function between points A and B depends upon the curve along which we integrate. (In general it does.) Consequently, an important ingredient in any complex integration is the *contour* that we follow during the integration.

The result of a line integral is a complex number or expression. Unlike its counterpart in real variables, there is no physical interpretation for this quantity, such as area under a curve. Generally, integration in the complex plane is an intermediate process with a physically realizable quantity occurring only after we take its real or imaginary part. For example, in potential fluid flow, the lift and drag are found by taking the real and imaginary part of a complex integral, respectively.

How do we compute $\int_C f(z)\,dz$? Let us deal with the definition; we illustrate the actual method by examples.

A popular method for evaluating complex line integrals consists of breaking everything up into real and imaginary parts. This reduces the integral to line integrals of real-valued functions, which we know how to handle. Thus, we write $f(z) = u(x, y) + iv(x, y)$ as usual, and because $z = x + iy$, formally $dz = dx + i\,dy$. Therefore,

$$\int_C f(z)\,dz = \int_C [u(x, y) + iv(x, y)][dx + i\,dy] \tag{10.4.1}$$

$$= \int_C u(x, y)\,dx - v(x, y)\,dy + i\int_C v(x, y)\,dx + u(x, y)\,dy. \tag{10.4.2}$$

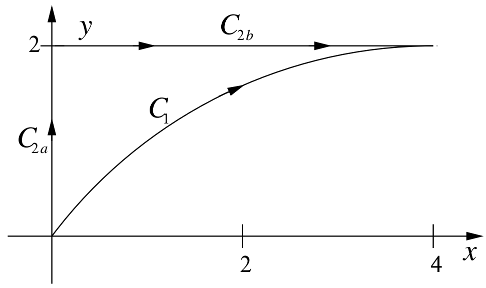


Figure 10.4.1: Contour used in Example 10.4.1.

The exact method used to evaluate Equation 10.4.2 depends upon the exact path specified.

From the definition of the line integral, we have the following self-evident properties:

$$\int_C f(z)\,dz = -\int_{C'} f(z)\,dz, \qquad (10.4.3)$$

where C' is the contour C taken in the opposite direction of C and

$$\int_{C_1+C_2} f(z)\,dz = \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz. \qquad (10.4.4)$$

• **Example 10.4.1**

Let us evaluate $\int_C z^* dz$ from $z = 0$ to $z = 4+2i$ along two different contours. The first consists of the parametric equation $z = t^2 + it$. The second consists of two "dog legs": the first leg runs along the imaginary axis from $z = 0$ to $z = 2i$ and then along a line parallel to the x-axis from $z = 2i$ to $z = 4 + 2i$. See Figure 10.4.1.

For the first case, the points $z = 0$ and $z = 4+2i$ on C_1 correspond to $t = 0$ and $t = 2$, respectively. Then the line integral equals

$$\int_{C_1} z^* dz = \int_0^2 (t^2 + it)^*\, d(t^2 + it) = \int_0^2 (2t^3 - it^2 + t)\, dt = 10 - \tfrac{8i}{3}. \qquad (10.4.5)$$

The line integral for the second contour C_2 equals

$$\int_{C_2} z^* dz = \int_{C_{2a}} z^* dz + \int_{C_{2b}} z^* dz, \qquad (10.4.6)$$

where C_{2a} denotes the integration from $z = 0$ to $z = 2i$ while C_{2b} denotes the integration from $z = 2i$ to $z = 4 + 2i$. For the first integral,

$$\int_{C_{2a}} z^* dz = \int_{C_{2a}} (x - iy)(dx + i\,dy) = \int_0^2 y\,dy = 2, \qquad (10.4.7)$$

because $x = 0$ and $dx = 0$ along C_{2a}. On the other hand, along C_{2b}, $y = 2$ and $dy = 0$ so that

$$\int_{C_{2b}} z^* dz = \int_{C_{2b}} (x - iy)(dx + i\,dy) = \int_0^4 x\,dx + i\int_0^4 -2\,dx = 8 - 8i. \qquad (10.4.8)$$

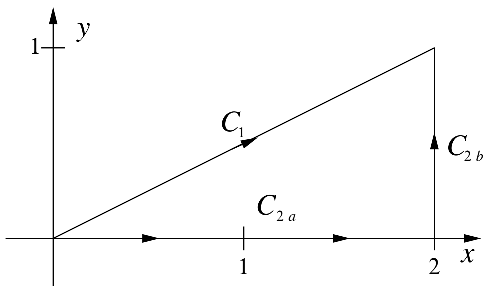


Figure 10.4.2: Contour used in Example 10.4.2.

Thus the value of the entire C_2 contour integral equals the sum of the two parts, or $10-8i$.

The point here is that integration along two different paths has given us different results even though we integrated from $z=0$ to $z=4+2i$ both times. This result foreshadows a general result that is extremely important. Because the integrand contains nonanalytic points along and inside the region enclosed by our two curves, as shown by the Cauchy-Riemann equations, the results depend upon the path taken. Since complex integrations often involve integrands that have nonanalytic points, many line integrations depend upon the contour taken. □

• **Example 10.4.2**

Let us integrate the *entire* function $f(z)=z^2$ along the two paths from $z=0$ to $z=2+i$ shown in Figure 10.4.2. For the first integration, $x=2y$, while along the second path we have two straight paths: $z=0$ to $z=2$ and $z=2$ to $z=2+i$.

For the first contour integration,

$$\int_{C_1} z^2 dz = \int_0^1 (2y+iy)^2(2\,dy+i\,dy) = \int_0^1 (3y^2+4y^2i)(2\,dy+i\,dy) \tag{10.4.9}$$

$$= \int_0^1 6y^2\,dy + 8y^2i\,dy + 3y^2i\,dy - 4y^2\,dy = \int_0^1 2y^2\,dy + 11y^2i\,dy \tag{10.4.10}$$

$$= \tfrac{2}{3}y^3\big|_0^1 + \tfrac{11}{3}iy^3\big|_0^1 = \tfrac{2}{3} + \tfrac{11i}{3}. \tag{10.4.11}$$

For our second integration,

$$\int_{C_2} z^2\,dz = \int_{C_{2a}} z^2\,dz + \int_{C_{2b}} z^2\,dz. \tag{10.4.12}$$

Along C_{2a} we find that $y=dy=0$ so that

$$\int_{C_{2a}} z^2\,dz = \int_0^2 x^2\,dx = \tfrac{1}{3}x^3\big|_0^2 = \tfrac{8}{3}, \tag{10.4.13}$$

and

$$\int_{C_{2b}} z^2\,dz = \int_0^1 (2+iy)^2 i\,dy = i\left(4y+2iy^2-\frac{y^3}{3}\right)\bigg|_0^1 = 4i-2-\tfrac{i}{3}, \tag{10.4.14}$$

because $x=2$ and $dx=0$. Consequently,

$$\int_{C_2} z^2\,dz = \frac{2}{3}+\frac{11i}{3}. \tag{10.4.15}$$

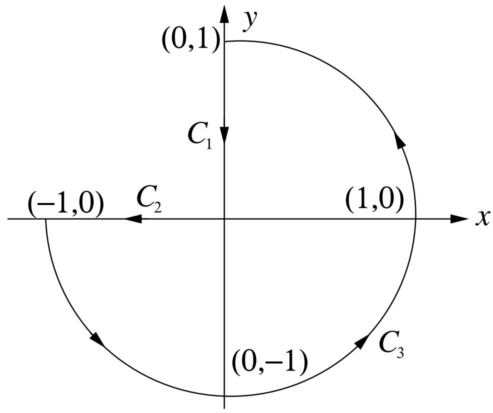


Figure 10.4.3: Contour used in Example 10.4.3.

In this problem we obtained the same results from two different contours of integration. Exploring other contours, we would find that the results are always the same; the integration is path-independent. But what makes these results path-independent while the integration in Example 10.4.1 was not? Perhaps it is the fact that the integrand is analytic everywhere on the complex plane and there are no nonanalytic points. We will explore this later. □

Finally, an important class of line integrals involves *closed contours*. We denote this special subclass of line integrals by placing a circle on the integral sign: $\oint$. Consider now the following examples:

• Example 10.4.3

Let us integrate $f(z) = z$ around the closed contour shown in Figure 10.4.3. From Figure 10.4.3,

$$\oint_C z\,dz = \int_{C_1} z\,dz + \int_{C_2} z\,dz + \int_{C_3} z\,dz. \tag{10.4.16}$$

Now

$$\int_{C_1} z\,dz = \int_1^0 iy\,(i\,dy) = -\int_1^0 y\,dy = -\left.\frac{y^2}{2}\right|_1^0 = \frac{1}{2}, \tag{10.4.17}$$

$$\int_{C_2} z\,dz = \int_0^{-1} x\,dx = \left.\frac{x^2}{2}\right|_0^{-1} = \frac{1}{2}, \tag{10.4.18}$$

and

$$\int_{C_3} z\,dz = \int_{-\pi}^{\pi/2} e^{\theta i} ie^{\theta i}\,d\theta = \left.\frac{e^{2\theta i}}{2}\right|_{-\pi}^{\pi/2} = -1, \tag{10.4.19}$$

where we used $z = e^{\theta i}$ around the portion of the unit circle. Therefore, the closed line integral equals zero. □

• Example 10.4.4

Let us integrate $f(z) = 1/(z - a)$ around any circle centered on $z = a$. The Cauchy-Riemann equations show that $f(z)$ is a meromorphic function. It is analytic everywhere except at the isolated singularity $z = a$.

If we introduce polar coordinates by letting $z - a = re^{\theta i}$ and $dz = ire^{\theta i} d\theta$,

$$\oint_C \frac{dz}{z-a} = \int_0^{2\pi} \frac{ire^{\theta i}}{re^{\theta i}}\, d\theta = i\int_0^{2\pi} d\theta = 2\pi i. \tag{10.4.20}$$

Note that the integrand becomes undefined at $z = a$. Furthermore, the answer is independent of the size of the circle. Our example suggests that when we have a closed contour integration it is the behavior of the function within the contour rather than the exact shape of the closed contour that is of importance. We will return to this point in later sections.

Problems

1. Evaluate $\oint_C (z^*)^2\, dz$ around the circle $|z| = 1$ taken in the counterclockwise direction.

2. Evaluate $\oint_C |z|^2\, dz$ around the square with vertices at (0,0), (1,0), (1,1), and (0,1) taken in the counterclockwise direction.

3. Evaluate $\int_C |z|\, dz$ along the right half of the circle $|z| = 1$ from $z = -i$ to $z = i$.

4. Evaluate $\int_C e^z\, dz$ along the line $y = x$ from $(-1, -1)$ to $(1, 1)$.

5. Evaluate $\int_C (z^*)^2\, dz$ along the line $y = x^2$ from $(0, 0)$ to $(1, 1)$.

6. Evaluate $\int_C z^{-1/2}\, dz$, where C is (a) the upper semicircle $|z| = 1$ and (b) the lower semicircle $|z| = 1$. If $z = re^{\theta i}$, restrict $-\pi < \theta < \pi$. Take both contours in the counterclockwise direction.

10.5 THE CAUCHY-GOURSAT THEOREM

In the previous section we showed how to evaluate line integrations by brute-force reduction to real-valued integrals. In general, this direct approach is quite difficult and we would like to apply some of the deeper properties of complex analysis to work smarter. In the remaining portions of this chapter we introduce several theorems that will do just that.

If we scan over the examples worked in the previous section, we see considerable differences when the function was analytic inside and on the contour and when it was not. We may formalize this anecdotal evidence into the following theorem:

Cauchy-Goursat theorem:[2] *Let $f(z)$ be analytic in a domain D and let C be a simple Jordan curve*[3] *inside D so that $f(z)$ is analytic on and inside of C. Then $\oint_C f(z)\, dz = 0$.*

Proof: Let C denote the contour around which we will integrate $w = f(z)$. We divide the region within C into a series of infinitesimal rectangles. See Figure 10.5.1. The integration

[2] Goursat, E., 1900: Sur la définition générale des fonctions analytiques, d'après Cauchy. *Trans. Am. Math. Soc.*, **1**, 14–16.

[3] A Jordan curve is a simply closed curve. It looks like a closed loop that does not cross itself. See Figure 10.5.2.

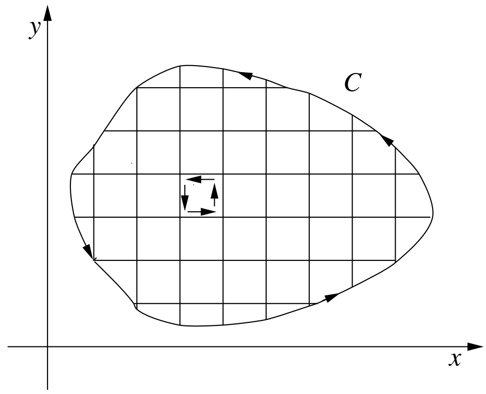


Figure 10.5.1: Diagram used in proving the Cauchy-Goursat theorem.

around each rectangle equals the product of the average value of w on each side and its length,

$$\left[w + \frac{\partial w}{\partial x}\frac{dx}{2}\right] dx + \left[w + \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial(iy)}\frac{d(iy)}{2}\right] d(iy)$$
$$+ \left[w + \frac{\partial w}{\partial x}\frac{dx}{2} + \frac{\partial w}{\partial(iy)}d(iy)\right](-dx) + \left[w + \frac{\partial w}{\partial(iy)}\frac{d(iy)}{2}\right] d(-iy)$$
$$= \left(\frac{\partial w}{\partial x} - \frac{\partial w}{i\partial y}\right)(i\,dx\,dy). \quad \mathbf{(10.5.1)}$$

Substituting $w = u + iv$ into Equation 10.5.1,

$$\frac{\partial w}{\partial x} - \frac{\partial w}{i\,\partial y} = \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + i\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right). \quad \mathbf{(10.5.2)}$$

Because the function is analytic, the right side of Equation 10.5.1 and Equation 10.5.2 equals zero. Thus, the integration around each of these rectangles also equals zero.

We note next that in integrating around adjoining rectangles we transverse each side in opposite directions, the net result being equivalent to integrating around the outer curve C. We therefore arrive at the result $\oint_C f(z)\,dz = 0$, where $f(z)$ is analytic within and on the closed contour.

□

The Cauchy-Goursat theorem has several useful implications. Suppose that we have a domain where $f(z)$ is analytic. Within this domain let us evaluate a line integral from point A to B along two different contours C_1 and C_2. Then, the integral around the closed contour formed by integrating along C_1 and then back along C_2, only in the opposite direction, is

$$\oint_C f(z)\,dz = \int_{C_1} f(z)\,dz - \int_{C_2} f(z)\,dz = 0 \quad \mathbf{(10.5.3)}$$

or

$$\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz. \quad \mathbf{(10.5.4)}$$

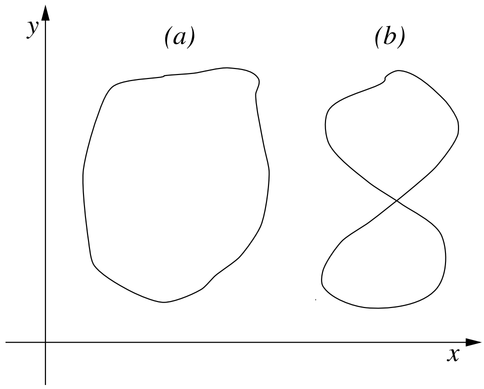


Figure 10.5.2: Examples of a (a) simply closed curve and (b) not simply closed curve.

Because C_1 and C_2 are completely arbitrary, we have the result that if, in a domain, $f(z)$ is analytic, the integral between any two points within the domain is *path independent.*

One obvious advantage of path independence is the ability to choose the contour so that the computations are made easier. This obvious choice immediately leads to

The principle of deformation of contours: *The value of a line integral of an analytic function around any simple closed contour remains unchanged if we deform the contour in such a manner that we do not pass over a nonanalytic point.* □

• Example 10.5.1

Let us integrate $f(z) = z^{-1}$ around the closed contour C in the counterclockwise direction. This contour consists of a square, centered on the origin, with vertices at $(1, 1)$, $(1, -1)$, $(-1, 1)$, and $(-1, -1)$.

The direct integration of $\oint_C z^{-1} dz$ around the original contour is very cumbersome. However, because the integrand is analytic everywhere except at the origin, we may deform the origin contour into a circle of radius r, centered on the origin. Then, $z = re^{\theta i}$ and $dz = rie^{\theta i} d\theta$ so that

$$\oint_C \frac{dz}{z} = \int_0^{2\pi} \frac{rie^{\theta i}}{re^{\theta i}}\, d\theta = i \int_0^{2\pi} d\theta = 2\pi i. \qquad \mathbf{(10.5.5)}$$

The point here is that no matter how bizarre the contour is, as long as it encircles the origin and is a simply closed contour, we can deform it into a circle and we get the same answer for the contour integral. This suggests that it is not the shape of the closed contour that makes the difference but whether we enclose any singularities (points where $f(z)$ becomes undefined) that matters. We shall return to this idea many times in the next few sections.□

Finally, suppose that we have a function $f(z)$ such that $f(z)$ is analytic in some domain. Furthermore, let us introduce the analytic function $F(z)$ such that $f(z) = F'(z)$. We would like to evaluate $\int_a^b f(z)\, dz$ in terms of $F(z)$.

We begin by noting that we can represent F, f as $F(z) = U + iV$ and $f(z) = u + iv$. From Example 10.3.28 we have that $u = U_x$ and $v = V_x$. Therefore,

$$\int_a^b f(z)\, dz = \int_a^b (u + iv)(dx + i\, dy) = \int_a^b U_x\, dx - V_x\, dy + i \int_a^b V_x\, dx + U_x\, dy \qquad \mathbf{(10.5.6)}$$

$$= \int_a^b U_x\,dx + U_y\,dy + i\int_a^b V_x\,dx + V_y\,dy = \int_a^b dU + i\int_a^b dV = F(b) - F(a) \tag{10.5.7}$$

or

$$\int_a^b f(z)\,dz = F(b) - F(a). \tag{10.5.8}$$

Equation 10.5.8 is the complex variable form of the fundamental theorem of calculus. Thus, if we can find the antiderivative of a function $f(z)$ that is analytic within a specific region, we can evaluate the integral by evaluating the antiderivative at the endpoints for any curves within that region.

• Example 10.5.2

Let us evaluate $\int_0^{\pi i} z\,\sin(z^2)\,dz$.

The integrand $f(z) = z\,\sin(z^2)$ is an entire function and its antiderivative equals $-\frac{1}{2}\cos(z^2)$. Therefore,

$$\int_0^{\pi i} z\,\sin(z^2)\,dz = -\tfrac{1}{2}\cos(z^2)\Big|_0^{\pi i} = \tfrac{1}{2}[\cos(0) - \cos(-\pi^2)] = \tfrac{1}{2}[1 - \cos(\pi^2)]. \tag{10.5.9}$$

Problems

For the following integrals, show that they are path independent and determine the value of the integral:

1. $\displaystyle\int_{1-\pi i}^{2+3\pi i} e^{-2z}\,dz$ 2. $\displaystyle\int_0^{2\pi} [e^z - \cos(z)]\,dz$ 3. $\displaystyle\int_0^{\pi} \sin^2(z)\,dz$ 4. $\displaystyle\int_{-i}^{2i} (z+1)\,dz$

10.6 CAUCHY'S INTEGRAL FORMULA

In the previous section, our examples suggested that the presence of a singularity within a contour really determines the value of a closed contour integral. Continuing with this idea, let us consider a class of closed contour integrals that explicitly contain a single singularity within the contour, namely $\oint_C g(z)\,dz$, where $g(z) = f(z)/(z - z_0)$, and $f(z)$ is analytic within and on the contour C. We closed the contour in the *positive sense* where the enclosed area lies to your left as you move along the contour.

We begin by examining a closed contour integral where the closed contour consists of the C_1, C_2, C_3, and C_4 as shown in Figure 10.6.1. The gap or cut between C_2 and C_4 is very small. Because $g(z)$ is analytic within and on the closed integral, we have that

$$\int_{C_1} \frac{f(z)}{z - z_0}\,dz + \int_{C_2} \frac{f(z)}{z - z_0}\,dz + \int_{C_3} \frac{f(z)}{z - z_0}\,dz + \int_{C_4} \frac{f(z)}{z - z_0}\,dz = 0. \tag{10.6.1}$$

It can be shown that the contribution to the integral from the path C_2 going into the singularity cancels the contribution from the path C_4 going away from the singularity as the gap between them vanishes. Because $f(z)$ is analytic at z_0, we can approximate its

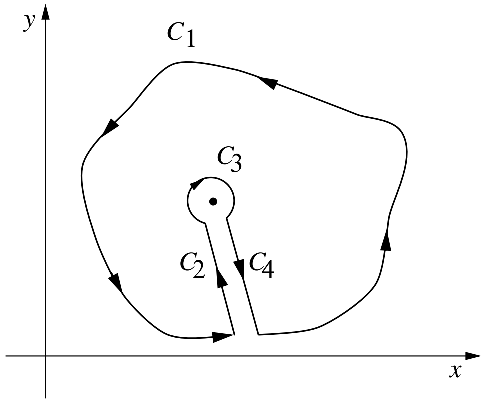


Figure 10.6.1: Diagram used to prove Cauchy's integral formula.

value on C_3 by $f(z) = f(z_0) + \delta(z)$, where δ is a small quantity. Substituting into Equation 10.6.1,

$$\oint_{C_1} \frac{f(z)}{z - z_0}\, dz = -f(z_0) \int_{C_3} \frac{1}{z - z_0}\, dz - \int_{C_3} \frac{\delta(z)}{z - z_0}\, dz. \tag{10.6.2}$$

Consequently, as the gap between C_2 and C_4 vanishes, the contour C_1 becomes the closed contour C so that Equation 10.6.2 may be written

$$\oint_{C} \frac{f(z)}{z - z_0}\, dz = 2\pi i f(z_0) + i \int_0^{2\pi} \delta\, d\theta, \tag{10.6.3}$$

where we set $z - z_0 = \epsilon e^{\theta i}$ and $dz = i\epsilon e^{\theta i} d\theta$.

Let M denote the value of the integral on the right side of Equation 10.6.3 and Δ equal the greatest value of the modulus of δ along the circle. Then

$$|M| < \int_0^{2\pi} |\delta|\, d\theta \le \int_0^{2\pi} \Delta\, d\theta = 2\pi\Delta. \tag{10.6.4}$$

As the radius of the circle diminishes to zero, Δ also diminishes to zero. Therefore, $|M|$, which is positive, becomes less than any finite quantity, however small, and M itself equals zero. Thus, we have that

$$\boxed{f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0}\, dz.} \tag{10.6.5}$$

This equation is *Cauchy's integral formula*. By taking n derivatives of Equation 10.6.5, we can extend Cauchy's integral formula[4] to

$$\boxed{f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz} \tag{10.6.6}$$

---

[4] See Carrier, G. F., M. Krook, and C. E. Pearson, 1966: *Functions of a Complex Variable: Theory and Technique*. McGraw-Hill, pp. 39–40 for the proof.

for $n = 1, 2, 3, \ldots$. For computing integrals, it is convenient to rewrite Equation 10.6.6 as

$$\oint_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz = \frac{2\pi i}{n!} f^{(n)}(z_0). \tag{10.6.7}$$

• **Example 10.6.1**

Let us find the value of the integral

$$\oint_C \frac{\cos(\pi z)}{(z-1)(z-2)}\, dz, \tag{10.6.8}$$

where C is the circle $|z| = 5$. Using partial fractions,

$$\frac{1}{(z-1)(z-2)} = \frac{1}{z-2} - \frac{1}{z-1}, \tag{10.6.9}$$

and

$$\oint_C \frac{\cos(\pi z)}{(z-1)(z-2)}\, dz = \oint_C \frac{\cos(\pi z)}{z-2}\, dz - \oint_C \frac{\cos(\pi z)}{z-1} dz. \tag{10.6.10}$$

By Cauchy's integral formula with $z_0 = 2$ and $z_0 = 1$,

$$\oint_C \frac{\cos(\pi z)}{z-2}\, dz = 2\pi i\ \cos(2\pi) = 2\pi i, \tag{10.6.11}$$

and

$$\oint_C \frac{\cos(\pi z)}{z-1}\, dz = 2\pi i \cos(\pi) = -2\pi i, \tag{10.6.12}$$

because $z_0 = 1$ and $z_0 = 2$ lie inside C and $\cos(\pi z)$ is analytic there. Thus the required integral has the value

$$\oint_C \frac{\cos(\pi z)}{(z-1)(z-2)}\, dz = 4\pi i. \tag{10.6.13}$$

□

• **Example 10.6.2**

Let us use Cauchy's integral formula to evaluate

$$I = \oint_{|z|=2} \frac{e^z}{(z-1)^2(z-3)}\, dz. \tag{10.6.14}$$

We need to convert Equation 10.6.14 into the form Equation 10.6.7. To do this, we rewrite Equation 10.6.14 as

$$\oint_{|z|=2} \frac{e^z}{(z-1)^2(z-3)}\, dz = \oint_{|z|=2} \frac{e^z/(z-3)}{(z-1)^2}\, dz. \tag{10.6.15}$$

Therefore, $f(z) = e^z/(z-3)$, $n = 1$, and $z_0 = 1$. The function $f(z)$ is analytic within the closed contour because the point $z = 3$ lies outside of the contour. Applying Cauchy's integral formula,

$$\oint_{|z|=2} \frac{e^z}{(z-1)^2(z-3)}\, dz = \frac{2\pi i}{1!} \frac{d}{dz}\left(\frac{e^z}{z-3}\right)\bigg|_{z=1} = 2\pi i \left[\frac{e^z}{z-3} - \frac{e^z}{(z-3)^2}\right]\bigg|_{z=1} = -\frac{3\pi i e}{2}. \tag{10.6.16}$$

Project: Computing Derivatives of Any Order of a Complex or Real Function

The most common technique for computing a derivative is finite differencing. Recently Mahajerin and Burgess[5] showed how Cauchy's integral formula can be used to compute the derivatives of any order of a complex or real function via numerical quadrature. In this project you will derive the algorithm, write code implementing it, and finally test it.

Step 1: Consider the complex function $f(z) = u + iv$, which is analytic inside the closed circular contour C of radius R centered at z_0. Using Cauchy's integral formula, show that

$$f^{(n)}(z_0) = \frac{n!}{2\pi R^n} \int_0^{2\pi} [u(x,y) + iv(x,y)][\cos(n\theta) - i\sin(n\theta)]\, d\theta,$$

where $x = x_0 + R\cos(\theta)$, and $y = y_0 + R\sin(\theta)$.

Step 2: Using five-point Gaussian quadrature, write code to implement the results from Step 1.

Step 3: Test out this scheme by finding the first, sixth, and eleventh derivative of $f(x) = 8x/(x^2+4)$ for $x = 2$. The exact answers are 0, 2.8125, and 1218.164, respectively. What is the maximum value of R? How does the accuracy vary with the number of subdivisions used in the numerical integration? Is the algorithm sensitive to the value of R and the number of subdivisions? For a fixed number of subdivisions, is there an optimal R?

Problems

Use Cauchy's integral formula to evaluate the following integrals. Assume all of the contours are in the positive sense.

1. $\displaystyle\oint_{|z|=1} \frac{\sin^6(z)}{z - \pi/6}\, dz$

2. $\displaystyle\oint_{|z|=1} \frac{\sin^6(z)}{(z - \pi/6)^3}\, dz$

3. $\displaystyle\oint_{|z|=1} \frac{1}{z(z^2+4)}\, dz$

4. $\displaystyle\oint_{|z|=1} \frac{\tan(z)}{z}\, dz$

5. $\displaystyle\oint_{|z-1|=1/2} \frac{1}{(z-1)(z-2)}\, dz$

6. $\displaystyle\oint_{|z|=5} \frac{\exp(z^2)}{z^3}\, dz$

7. $\displaystyle\oint_{|z-1|=1} \frac{z^2+1}{z^2-1}\, dz$

8. $\displaystyle\oint_{|z|=2} \frac{z^2}{(z-1)^4}\, dz$

9. $\displaystyle\oint_{|z|=2} \frac{z^3}{(z+i)^3}\, dz$

10. $\displaystyle\oint_{|z|=1} \frac{\cos(z)}{z^{2n+1}}\, dz$

10.7 TAYLOR AND LAURENT EXPANSIONS AND SINGULARITIES

In the previous section we showed what a crucial role singularities play in complex integration. Before we can find the most general way of computing a closed complex integral, our understanding of singularities must deepen. For this, we employ power series.

---

[5] Mahajerin, E., and G. Burgess, 1993: An algorithm for computing derivatives of any order of a complex or real function. *Computers & Struct.*, **49**, 385–387.

One reason why power series are so important is their ability to provide locally a general representation of a function even when its arguments are complex. For example, when we were introduced to trigonometric functions in high school, it was in the context of a right triangle and a real angle. However, when the argument becomes complex, this geometrical description disappears and power series provide a formalism for defining the trigonometric functions, regardless of the nature of the argument.

Let us begin our analysis by considering the complex function $f(z)$, which is analytic everywhere on the boundary, and the interior of a circle whose center is at $z = z_0$. Then, if z denotes any point within the circle, we have from Cauchy's integral formula that

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{\zeta - z}\, d\zeta = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{\zeta - z_0} \left[\frac{1}{1-(z-z_0)/(\zeta - z_0)}\right] d\zeta, \qquad \textbf{(10.7.1)}$$

where C denotes the closed contour. Expanding the bracketed term as a geometric series, we find that

$$f(z) = \frac{1}{2\pi i}\left[\oint_C \frac{f(\zeta)}{\zeta - z_0}\, d\zeta + (z-z_0) \oint_C \frac{f(\zeta)}{(\zeta - z_0)^2}\, d\zeta + \cdots + (z-z_0)^n \oint_C \frac{f(\zeta)}{(\zeta - z_0)^{n+1}}\, d\zeta + \cdots\right]. \qquad \textbf{(10.7.2)}$$

Applying Cauchy's integral formula to each integral in Equation 10.7.2, we finally obtain

$$f(z) = f(z_0) + \frac{(z-z_0)}{1!} f'(z_0) + \cdots + \frac{(z-z_0)^n}{n!} f^{(n)}(z_0) + \cdots \qquad \textbf{(10.7.3)}$$

or the familiar formula for a Taylor expansion. Consequently, *we can expand any analytic function into a Taylor series.* Interestingly, the radius of convergence[6] of this series may be shown to be the distance between z_0 and the nearest nonanalytic point of $f(z)$.

• Example 10.7.1

Let us find the expansion of $f(z) = \sin(z)$ about the point $z_0 = 0$.

Because $f(z)$ is an entire function, we can construct a Taylor expansion anywhere on the complex plane. For $z_0 = 0$,

$$f(z) = f(0) + \tfrac{1}{1!} f'(0) z + \tfrac{1}{2!} f''(0) z^2 + \tfrac{1}{3!} f'''(0) z^3 + \cdots. \qquad \textbf{(10.7.4)}$$

Because $f(0) = 0$, $f'(0) = 1$, $f''(0) = 0$, $f'''(0) = -1$ and so forth,

$$f(z) = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots. \qquad \textbf{(10.7.5)}$$

Because $\sin(z)$ is an entire function, the radius of convergence is $|z - 0| < \infty$, i.e., all z. □

• Example 10.7.2

Let us find the expansion of $f(z) = 1/(1-z)$ about the point $z_0 = 0$.

From the formula for a Taylor expansion,

$$f(z) = f(0) + \tfrac{1}{1!} f'(0) z + \tfrac{1}{2!} f''(0) z^2 + \tfrac{1}{3!} f'''(0) z^3 + \cdots. \qquad \textbf{(10.7.6)}$$

[6] A positive number h such that the series diverges for $|z - z_0| > h$ but converges absolutely for $|z - z_0| < h$.

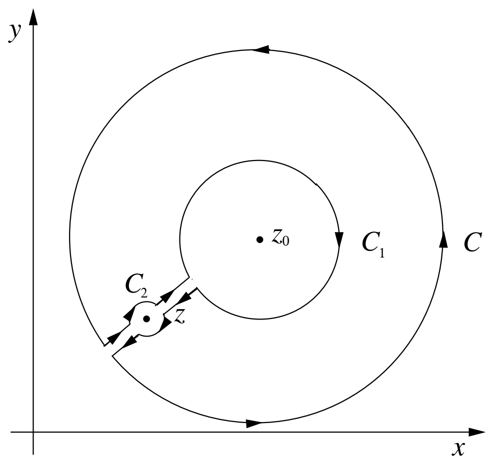


Figure 10.7.1: Contour used in deriving the Laurent expansion.

Because $f^{(n)}(0) = n!$, we find that

$$f(z) = 1 + z + z^2 + z^3 + z^4 + \cdots = \frac{1}{1-z}. \qquad (\mathbf{10.7.7})$$

Equation 10.7.7 is the familiar result for a geometric series. Because the only nonanalytic point is at $z = 1$, the radius of convergence is $|z-0| < 1$, the unit circle centered at $z = 0$.□

Consider now the situation where we draw two concentric circles about some arbitrary point z_0; we denote the outer circle by C while we denote the inner circle by C_1. See Figure 10.7.1. Let us assume that $f(z)$ is analytic inside the annulus between the two circles. Outside of this area, the function may or may not be analytic. Within the annulus we pick a point z and construct a small circle around it, denoting the circle by C_2. As the gap or *cut* in the annulus becomes infinitesimally small, the line integrals that connect the circle C_2 to C_1 and C sum to zero, leaving

$$\oint_C \frac{f(\zeta)}{\zeta - z}\, d\zeta = \oint_{C_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta + \oint_{C_2} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \qquad (\mathbf{10.7.8})$$

Because $f(\zeta)$ is analytic everywhere within C_2,

$$2\pi i f(z) = \oint_{C_2} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \qquad (\mathbf{10.7.9})$$

Using the relationship:

$$\oint_{C_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta = -\oint_{C_1} \frac{f(\zeta)}{z - \zeta}\, d\zeta, \qquad (\mathbf{10.7.10})$$

Equation 10.7.8 becomes

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{\zeta - z}\, d\zeta + \frac{1}{2\pi i} \oint_{C_1} \frac{f(\zeta)}{z - \zeta}\, d\zeta. \qquad (\mathbf{10.7.11})$$

Now,

$$\frac{1}{\zeta - z} = \frac{1}{\zeta - z_0 - z + z_0} = \frac{1}{\zeta - z_0} \frac{1}{1 - (z - z_0)/(\zeta - z_0)} \qquad (\mathbf{10.7.12})$$

$$= \frac{1}{\zeta - z_0}\left[1 + \left(\frac{z - z_0}{\zeta - z_0}\right) + \left(\frac{z - z_0}{\zeta - z_0}\right)^2 + \cdots + \left(\frac{z - z_0}{\zeta - z_0}\right)^n + \cdots\right], \qquad (\mathbf{10.7.13})$$

where $|z - z_0|/|\zeta - z_0| < 1$ and

$$\frac{1}{z-\zeta} = \frac{1}{z - z_0 - \zeta + z_0} = \frac{1}{z-z_0}\frac{1}{1-(\zeta-z_0)/(z-z_0)} \tag{10.7.14}$$

$$= \frac{1}{z-z_0}\left[1 + \left(\frac{\zeta-z_0}{z-z_0}\right) + \left(\frac{\zeta-z_0}{z-z_0}\right)^2 + \cdots + \left(\frac{\zeta-z_0}{z-z_0}\right)^n + \cdots\right], \tag{10.7.15}$$

where $|\zeta - z_0|/|z - z_0| < 1$. Upon substituting these expressions into Equation 10.7.11,

$$\begin{aligned} f(z) = &\left[\frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta - z_0}\,d\zeta + \frac{z-z_0}{2\pi i}\oint_C \frac{f(\zeta)}{(\zeta-z_0)^2}\,d\zeta + \cdots \right.\\ &\left. + \frac{(z-z_0)^n}{2\pi i}\oint_C \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}\,d\zeta + \cdots\right] \\ &+ \left[\frac{1}{z-z_0}\frac{1}{2\pi i}\oint_{C_1} f(\zeta)\,d\zeta + \frac{1}{(z-z_0)^2}\frac{1}{2\pi i}\oint_{C_1} f(\zeta)(\zeta-z_0)\,d\zeta + \cdots\right.\\ &\left. + \frac{1}{(z-z_0)^n}\frac{1}{2\pi i}\oint_{C_1} f(\zeta)(\zeta-z_0)^{n-1}\,d\zeta + \cdots\right] \end{aligned} \tag{10.7.16}$$

or

$$f(z) = \frac{a_1}{z-z_0} + \frac{a_2}{(z-z_0)^2} + \cdots + \frac{a_n}{(z-z_0)^n} + \cdots + b_0 + b_1(z-z_0) + \cdots + b_n(z-z_0)^n + \cdots. \tag{10.7.17}$$

Equation 10.7.17 is a *Laurent expansion.*[7] If $f(z)$ is analytic at z_0, then $a_1 = a_2 = \cdots = a_n = \cdots = 0$ and the Laurent expansion reduces to a Taylor expansion. If z_0 is a singularity of $f(z)$, then the Laurent expansion includes both positive and *negative* powers. The coefficient of the $(z - z_0)^{-1}$ term, a_1, is the *residue*, for reasons that will appear in the next section.

Unlike the Taylor series, a Laurent series provides no straightforward method for obtaining the coefficients. For the remaining portions of this section we illustrate their construction. These techniques include replacing a function by its appropriate power series, the use of geometric series to expand the denominator, and the use of algebraic tricks to assist in applying the first two methods.

• Example 10.7.3

Laurent expansions provide a formalism for the classification of singularities of a function. *Isolated singularities* fall into three types; they are

• *Essential Singularity*: Consider the function $f(z) = \cos(1/z)$. Using the expansion for cosine,

$$\cos\left(\frac{1}{z}\right) = 1 - \frac{1}{2!z^2} + \frac{1}{4!z^4} - \frac{1}{6!z^6} + \cdots \tag{10.7.18}$$

for $0 < |z| < \infty$. Note that this series never truncates in the inverse powers of z. Essential singularities have Laurent expansions, which have an infinite number of inverse powers of $z - z_0$. The value of the residue for this essential singularity at $z = 0$ is zero.

---

[7] Laurent, M., 1843: Extension du théorème de M. Cauchy relatif à la convergence du développement d'une fonction suivant les puissances ascendantes de la variable x. *C. R. l'Acad. Sci.*, **17**, 938–942.

• *Removable Singularity*: Consider the function $f(z) = \sin(z)/z$. This function has a singularity at $z = 0$. Upon applying the expansion for sine,

$$\frac{\sin(z)}{z} = \frac{1}{z}\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \frac{z^9}{9!} - \ldots\right) = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \frac{z^6}{7!} + \frac{z^8}{9!} - \ldots \quad \textbf{(10.7.19)}$$

for all z, if the division is permissible. We made $f(z)$ analytic by defining it by Equation 10.7.19 and, in the process, removed the singularity. The residue for a removable singularity always equals zero.

• *Pole of order n*: Consider the function

$$f(z) = \frac{1}{(z-1)^3(z+1)}. \quad \textbf{(10.7.20)}$$

This function has two singularities: one at $z = 1$ and the other at $z = -1$. We shall only consider the case $z = 1$. After a little algebra,

$$f(z) = \frac{1}{(z-1)^3}\,\frac{1}{2+(z-1)} = \frac{1}{2}\frac{1}{(z-1)^3}\,\frac{1}{1+(z-1)/2} \quad \textbf{(10.7.21)}$$

$$= \frac{1}{2}\frac{1}{(z-1)^3}\left[1 - \frac{z-1}{2} + \frac{(z-1)^2}{4} - \frac{(z-1)^3}{8} + \cdots\right] \quad \textbf{(10.7.22)}$$

$$= \frac{1}{2(z-1)^3} - \frac{1}{4(z-1)^2} + \frac{1}{8(z-1)} - \frac{1}{16} + \cdots \quad \textbf{(10.7.23)}$$

for $0 < |z-1| < 2$. Because the largest inverse (negative) power is three, the singularity at $z = 1$ is a third-order pole; the value of the residue is 1/8. Generally, we refer to a first-order pole as a *simple* pole. □

• **Example 10.7.4**

Let us find the Laurent expansion for

$$f(z) = \frac{z}{(z-1)(z-3)} \quad \textbf{(10.7.24)}$$

about the point $z = 1$.

We begin by rewriting $f(z)$ as

$$f(z) = \frac{1+(z-1)}{(z-1)[-2+(z-1)]} = -\frac{1}{2}\,\frac{1+(z-1)}{(z-1)[1-\frac{1}{2}(z-1)]} \quad \textbf{(10.7.25)}$$

$$= -\frac{1}{2}\,\frac{1+(z-1)}{(z-1)}[1 + \tfrac{1}{2}(z-1) + \tfrac{1}{4}(z-1)^2 + \cdots] \quad \textbf{(10.7.26)}$$

$$= -\frac{1}{2}\,\frac{1}{z-1} - \frac{3}{4} - \frac{3}{8}(z-1) - \frac{3}{16}(z-1)^2 - \cdots \quad \textbf{(10.7.27)}$$

provided $0 < |z-1| < 2$. Therefore we have a simple pole at $z = 1$ and the value of the residue is $-1/2$. A similar procedure would yield the Laurent expansion about $z = 3$. □

• Example 10.7.5

Let us find the Laurent expansion for

$$f(z) = \frac{z^n + z^{-n}}{z^2 - 2z\cosh(\alpha) + 1}, \quad \alpha > 0, \quad n \geq 0, \tag{10.7.28}$$

about the point $z = 0$.

We begin by rewriting $f(z)$ as

$$f(z) = \frac{z^n + z^{-n}}{(z - e^\alpha)(z - e^{-\alpha})} = \frac{1}{2\sinh(\alpha)}\left(\frac{z^n + z^{-n}}{z - e^\alpha} - \frac{z^n + z^{-n}}{z - e^{-\alpha}}\right). \tag{10.7.29}$$

Because

$$\frac{1}{z - e^\alpha} = -\frac{e^{-\alpha}}{1 - ze^{-\alpha}} = -e^{-\alpha}\left(1 + ze^{-\alpha} + z^2e^{-2\alpha} + \cdots\right) \tag{10.7.30}$$

if $|z| < e^\alpha$ and

$$\frac{1}{z - e^{-\alpha}} = -\frac{e^\alpha}{1 - ze^\alpha} = -e^\alpha\left(1 + ze^\alpha + z^2e^{2\alpha} + \cdots\right) \tag{10.7.31}$$

if $|z| < e^{-\alpha}$,

$$\begin{aligned} f(z) = {} & \frac{e^\alpha}{2\sinh(\alpha)}\left(z^n + z^{n+1}e^\alpha + z^{n+2}e^{2\alpha} + \cdots + z^{-n} + z^{1-n}e^\alpha + z^{2-n}e^{2\alpha} + \cdots\right) \\ & - \frac{e^{-\alpha}}{2\sinh(\alpha)}\left(z^n + z^{n+1}e^{-\alpha} + z^{n+2}e^{-2\alpha} + \cdots + z^{-n} + z^{1-n}e^{-\alpha} + z^{2-n}e^{-2\alpha} + \cdots\right), \end{aligned} \tag{10.7.32}$$

if $|z| < e^{-\alpha}$. Clearly we have an nth-order pole at $z = 0$. The residue, the coefficient of all of the z^{-1} terms in Equation 10.7.32, is found directly and equals

$$\text{Res}[f(z); 0] = \frac{\sinh(n\alpha)}{\sinh(\alpha)}. \tag{10.7.33}$$

□

For complicated complex functions, it is very difficult to determine the nature of the singularities by finding the complete Laurent expansion, and we must try another method. We shall call it "a poor man's Laurent expansion." The idea behind this method is the fact that we generally need only the first few terms of the Laurent expansion to discover its nature. Consequently, we compute these terms through the application of power series where we retain only the leading terms. Consider the following example.

• Example 10.7.6

Let us discover the nature of the singularity at $z = 0$ of the function

$$f(z) = \frac{e^{tz}}{z\sinh(az)}, \tag{10.7.34}$$

where a and t are real.

We begin by replacing the exponential and hyperbolic sine by their Taylor expansion about $z = 0$. Then

$$f(z) = \frac{1 + tz + t^2z^2/2 + \cdots}{z(az + a^3z^3/6 + \cdots)}. \tag{10.7.35}$$

Factoring out az in the denominator,

$$f(z) = \frac{1 + tz + t^2z^2/2 + \cdots}{az^2(1 + a^2z^2/6 + \cdots)}. \tag{10.7.36}$$

Within the parentheses all of the terms except the leading one are small. Therefore, by long division, we formally have that

$$f(z) = \frac{1}{az^2}(1 + tz + t^2z^2/2 + \cdots)(1 - a^2z^2/6 + \cdots) \tag{10.7.37}$$

$$= \frac{1}{az^2}(1 + tz + t^2z^2/2 - a^2z^2/6 + \cdots) = \frac{1}{az^2} + \frac{t}{az} + \frac{3t^2 - a^2}{6a} + \cdots. \tag{10.7.38}$$

Thus, we have a second-order pole at $z = 0$ and the residue equals t/a.

Problems

1. Find the Taylor expansion of $f(z) = (1-z)^{-2}$ about the point $z = 0$.

2. Find the Taylor expansion of $f(z) = (z-1)e^z$ about the point $z = 1$. (Hint: Don't find the expansion by taking derivatives.)

By constructing a Laurent expansion, describe the type of singularity and give the residue at z_0 for each of the following functions:

3. $f(z) = z^{10}e^{-1/z}; \quad z_0 = 0$

4. $f(z) = z^{-3}\sin^2(z); \quad z_0 = 0$

5. $f(z) = \dfrac{\cosh(z) - 1}{z^2}; \quad z_0 = 0$

6. $f(z) = \dfrac{z}{(z+2)^2}; \quad z_0 = -2$

7. $f(z) = \dfrac{e^z + 1}{e^{-z} - 1}; \quad z_0 = 0$

8. $f(z) = \dfrac{e^{iz}}{z^2 + b^2}; \quad z_0 = bi$

9. $f(z) = \dfrac{1}{z(z-2)}; \quad z_0 = 2$

10. $f(z) = \dfrac{\exp(z^2)}{z^4}; \quad z_0 = 0$

10.8 THEORY OF RESIDUES

Having shown that around any singularity we may construct a Laurent expansion, we now use this result in the integration of closed complex integrals. Consider a closed contour in which the function $f(z)$ has a number of isolated singularities. As we did in the case of Cauchy's integral formula, we introduce a new contour C' that excludes all of the singularities because they are isolated. See Figure 10.8.1. Therefore,

$$\oint_C f(z)\,dz - \oint_{C_1} f(z)\,dz - \cdots - \oint_{C_n} f(z)\,dz = \oint_{C'} f(z)\,dz = 0. \tag{10.8.1}$$

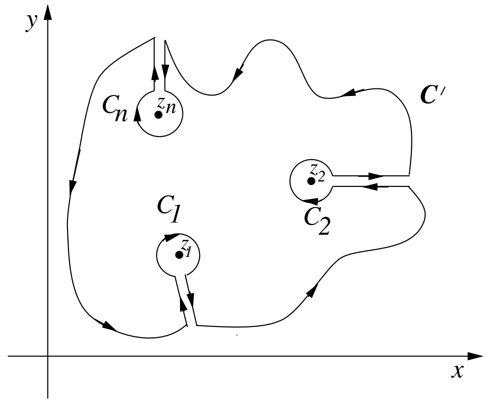


Figure 10.8.1: Contour used in deriving the residue theorem.

Consider now the mth integral, where $1 \le m \le n$. Constructing a Laurent expansion for the function $f(z)$ at the isolated singularity $z = z_m$, this integral equals

$$\oint_{C_m} f(z)\,dz = \sum_{k=1}^{\infty} a_k \oint_{C_m} \frac{1}{(z-z_m)^k}\,dz + \sum_{k=0}^{\infty} b_k \oint_{C_m} (z-z_m)^k\,dz. \qquad \textbf{(10.8.2)}$$

Because $(z-z_m)^k$ is an entire function if $k \ge 0$, the integrals equal zero for each term in the second summation. We use Cauchy's integral formula to evaluate the remaining terms. The analytic function in the numerator is 1. Because $d^{k-1}(1)/dz^{k-1} = 0$ if $k > 1$, all of the terms vanish except for $k = 1$. In that case, the integral equals $2\pi i a_1$, where a_1 is the value of the residue for that particular singularity. Applying this approach to each of the singularities, we obtain

Cauchy's residue theorem:[8] *If $f(z)$ is analytic inside and on a closed contour C (taken in the positive sense) except at points $z_1, z_2, \ldots, z_n$ where $f(z)$ has singularities, then*

$$\oint_C f(z)\,dz = 2\pi i \sum_{j=1}^{n} \operatorname{Res}[f(z); z_j], \qquad \textbf{(10.8.3)}$$

where $\operatorname{Res}[f(z); z_j]$ *denotes the residue of the jth isolated singularity of $f(z)$ located at $z = z_j$.* □

• Example 10.8.1

Let us compute $\oint_{|z|=2} z^2/(z+1)\,dz$ by the residue theorem, assuming that we take the contour in the positive sense.

Because the contour is a circle of radius 2, centered on the origin, the singularity at $z = -1$ lies within the contour. If the singularity were not inside the contour, then the

[8] See Mitrinović, D. S., and J. D. Kečkić, 1984: *The Cauchy Method of Residues: Theory and Applications*. D. Reidel Publishing, 361 pp. Section 10.3 gives the historical development of the residue theorem.

integrand would have been analytic inside and on the contour C. In this case, the answer would then be zero by the Cauchy-Goursat theorem.

Returning to the original problem, we construct the Laurent expansion for the integrand around the point $z = 1$ by noting that

$$\frac{z^2}{z+1} = \frac{[(z+1)-1]^2}{z+1} = \frac{1}{z+1} - 2 + (z+1). \tag{10.8.4}$$

The singularity at $z = -1$ is a simple pole and by inspection the value of the residue equals 1. Therefore,

$$\oint_{|z|=2} \frac{z^2}{z+1}\, dz = 2\pi i. \tag{10.8.5}$$

□

As it presently stands, it would appear that we must always construct a Laurent expansion for each singularity if we wish to use the residue theorem. This becomes increasingly difficult as the structure of the integrand becomes more complicated. In the following paragraphs we show several techniques that avoid this problem in practice.

We begin by noting that many functions which we will encounter consist of the ratio of two *polynomials*, i.e., rational functions: $f(z) = g(z)/h(z)$. Generally, we can write $h(z)$ as $(z-z_1)^{m_1}(z-z_2)^{m_2}\cdots$. Here we assumed that we divided out any common factors between $g(z)$ and $h(z)$ so that $g(z)$ does not vanish at z_1, z_2, ... Clearly $z_1, z_2, \ldots$, are singularities of $f(z)$. Further analysis shows that the nature of the singularities are a pole of order m_1 at $z = z_1$, a pole of order m_2 at $z = z_2$, and so forth.

Having found the nature and location of the singularity, we compute the residue as follows. Suppose that we have a pole of order n. Then we know that its Laurent expansion is

$$f(z) = \frac{a_n}{(z-z_0)^n} + \frac{a_{n-1}}{(z-z_0)^{n-1}} + \cdots + b_0 + b_1(z-z_0) + \cdots. \tag{10.8.6}$$

Multiplying both sides of Equation 10.8.6 by $(z-z_0)^n$,

$$F(z) = (z-z_0)^n f(z) = a_n + a_{n-1}(z-z_0) + \cdots + b_0(z-z_0)^n + b_1(z-z_0)^{n+1} + \cdots. \tag{10.8.7}$$

Because $F(z)$ is analytic at $z = z_0$, it has the Taylor expansion

$$F(z) = F(z_0) + F'(z_0)(z-z_0) + \cdots + \frac{F^{(n-1)}(z_0)}{(n-1)!}(z-z_0)^{n-1} + \cdots. \tag{10.8.8}$$

Matching powers of $z - z_0$ in Equation 10.8.7 and Equation 10.8.8, the residue equals

$$\text{Res}[f(z); z_0] = a_1 = \frac{F^{(n-1)}(z_0)}{(n-1)!}. \tag{10.8.9}$$

Substituting in $F(z) = (z-z_0)^n f(z)$, we can compute the residue of a pole of order n by

$$\boxed{\text{Res}[f(z); z_j] = \frac{1}{(n-1)!} \lim_{z\to z_j} \frac{d^{n-1}}{dz^{n-1}}\left[(z-z_j)^n f(z)\right].} \tag{10.8.10}$$

For a simple pole, Equation 10.8.10 simplifies to

$$\text{Res}[f(z); z_j] = \lim_{z \to z_j} (z - z_j) f(z). \tag{10.8.11}$$

Quite often, $f(z) = p(z)/q(z)$. From l'Hôspital's rule, it follows that Equation 10.8.11 becomes

$$\text{Res}[f(z); z_j] = \frac{p(z_j)}{q'(z_j)}. \tag{10.8.12}$$

Recall that these formulas work only for finite-order poles. For an essential singularity we must compute the residue from its Laurent expansion; however, essential singularities are very rare in applications.

• **Example 10.8.2**

Let us evaluate

$$\oint_C \frac{e^{iz}}{z^2 + a^2}\, dz, \tag{10.8.13}$$

where C is any contour that includes both poles at $z = \pm ai$ and is in the positive sense.

From Cauchy's residue theorem,

$$\oint_C \frac{e^{iz}}{z^2 + a^2}\, dz = 2\pi i \left[\text{Res}\left(\frac{e^{iz}}{z^2 + a^2}; ai \right) + \text{Res}\left(\frac{e^{iz}}{z^2 + a^2}; -ai \right) \right]. \tag{10.8.14}$$

The singularities at $z = \pm ai$ are simple poles. The corresponding residues are

$$\text{Res}\left(\frac{e^{iz}}{z^2 + a^2}; ai \right) = \lim_{z \to ai} (z - ai) \frac{e^{iz}}{(z - ai)(z + ai)} = \frac{e^{-a}}{2ia} \tag{10.8.15}$$

and

$$\text{Res}\left(\frac{e^{iz}}{z^2 + a^2}; -ai \right) = \lim_{z \to -ai} (z + ai) \frac{e^{iz}}{(z - ai)(z + ai)} = -\frac{e^{a}}{2ia}. \tag{10.8.16}$$

Consequently,

$$\oint_C \frac{e^{iz}}{z^2 + a^2}\, dz = -\frac{2\pi}{2a} \left(e^{a} - e^{-a} \right) = -\frac{2\pi}{a} \sinh(a). \tag{10.8.17}$$

□

• **Example 10.8.3**

Let us evaluate

$$\frac{1}{2\pi i} \oint_C \frac{e^{tz}}{z^2(z^2 + 2z + 2)}\, dz, \tag{10.8.18}$$

where C includes all of the singularities and is in the positive sense.

The integrand has a second-order pole at $z = 0$ and two simple poles at $z = -1 \pm i$, which are the roots of $z^2 + 2z + 2 = 0$. Therefore, the residue at $z = 0$ is

$$\operatorname{Res}\left[\frac{e^{tz}}{z^2(z^2+2z+2)}; 0\right] = \lim_{z\to 0} \frac{1}{1!}\frac{d}{dz}\left\{(z-0)^2\left[\frac{e^{tz}}{z^2(z^2+2z+2)}\right]\right\} \tag{10.8.19}$$

$$= \lim_{z\to 0}\left[\frac{te^{tz}}{z^2+2z+2} - \frac{(2z+2)e^{tz}}{(z^2+2z+2)^2}\right] = \frac{t-1}{2}. \tag{10.8.20}$$

The residue at $z = -1 + i$ is

$$\operatorname{Res}\left[\frac{e^{tz}}{z^2(z^2+2z+2)}; -1+i\right] = \lim_{z\to -1+i}[z-(-1+i)]\frac{e^{tz}}{z^2(z^2+2z+2)} \tag{10.8.21}$$

$$= \left(\lim_{z\to -1+i}\frac{e^{tz}}{z^2}\right)\left(\lim_{z\to -1+i}\frac{z+1-i}{z^2+2z+2}\right) \tag{10.8.22}$$

$$= \frac{\exp[(-1+i)t]}{2i(-1+i)^2} = \frac{\exp[(-1+i)t]}{4}. \tag{10.8.23}$$

Similarly, the residue at $z = -1 - i$ is

$$\operatorname{Res}\left[\frac{e^{tz}}{z^2(z^2+2z+2)}; -1-i\right] = \lim_{z\to -1-i}[z-(-1-i)]\frac{e^{tz}}{z^2(z^2+2z+2)} \tag{10.8.24}$$

$$= \left(\lim_{z\to -1-i}\frac{e^{tz}}{z^2}\right)\left(\lim_{z\to -1-i}\frac{z+1+i}{z^2+2z+2}\right) \tag{10.8.25}$$

$$= \frac{\exp[(-1-i)t]}{(-2i)(-1-i)^2} = \frac{\exp[(-1-i)t]}{4}. \tag{10.8.26}$$

Then by the residue theorem,

$$\frac{1}{2\pi i}\oint_C \frac{e^{tz}}{z^2(z^2+2z+2)}\,dz = \operatorname{Res}\left[\frac{e^{tz}}{z^2(z^2+2z+2)}; 0\right] + \operatorname{Res}\left[\frac{e^{tz}}{z^2(z^2+2z+2)}; -1+i\right]$$

$$+ \operatorname{Res}\left[\frac{e^{tz}}{z^2(z^2+2z+2)}; -1-i\right] \tag{10.8.27}$$

$$= \frac{t-1}{2} + \frac{\exp[(-1+i)t]}{4} + \frac{\exp[(-1-i)t]}{4} \tag{10.8.28}$$

$$= \tfrac{1}{2}\left[t - 1 + e^{-t}\cos(t)\right]. \tag{10.8.29}$$

Problems

Assuming that all of the following closed contours are in the positive sense, use the residue theorem to evaluate the following integrals:

1. $\displaystyle\oint_{|z|=1} \frac{z+1}{z^4-2z^3}\,dz$ 2. $\displaystyle\oint_{|z|=1} \frac{(z+4)^3}{z^4+5z^3+6z^2}\,dz$ 3. $\displaystyle\oint_{|z|=1} \frac{1}{1-e^z}\,dz$

4. $\displaystyle\oint_{|z|=2} \frac{z^2-4}{(z-1)^4}\,dz$ 5. $\displaystyle\oint_{|z|=2} \frac{z^3}{z^4-1}\,dz$ 6. $\displaystyle\oint_{|z|=1} z^n e^{2/z}\,dz, \quad n > 0$

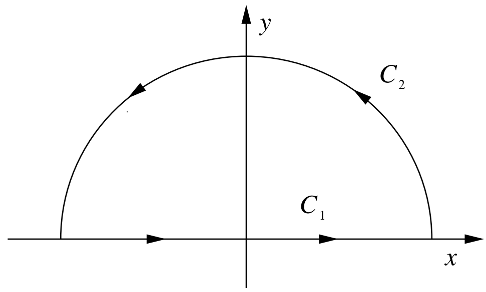


Figure 10.9.1: Contour used in evaluating the integral, Equation 10.9.1.

7. $\oint_{|z|=1} e^{1/z}\cos(1/z)\,dz$ 8. $\oint_{|z|=2} \frac{2+4\cos(\pi z)}{z(z-1)^2}\,dz$

10.9 EVALUATION OF REAL DEFINITE INTEGRALS

One of the important applications of the theory of residues consists in the evaluation of certain types of real definite integrals. Similar techniques apply when the integrand contains a sine or cosine. See Section 11.4.

• **Example 10.9.1**

Let us evaluate the integral

$$\int_0^\infty \frac{dx}{x^2+1} = \frac{1}{2}\int_{-\infty}^\infty \frac{dx}{x^2+1}. \tag{10.9.1}$$

This integration occurs along the real axis. In terms of complex variables, we can rewrite Equation 10.9.1 as

$$\int_0^\infty \frac{dx}{x^2+1} = \frac{1}{2}\int_{C_1} \frac{dz}{z^2+1}, \tag{10.9.2}$$

where the contour C_1 is the line $\Im(z)=0$. However, the use of the residue theorem requires an integration along a closed contour. Let us choose the one pictured in Figure 10.9.1. Then

$$\oint_C \frac{dz}{z^2+1} = \int_{C_1} \frac{dz}{z^2+1} + \int_{C_2} \frac{dz}{z^2+1}, \tag{10.9.3}$$

where C denotes the complete closed contour and C_2 denotes the integration path along a semicircle at infinity. Clearly we want the second integral on the right side of Equation 10.9.3 to vanish; otherwise, our choice of the contour C_2 is poor. Because $z = Re^{\theta i}$ and $dz = iRe^{\theta i}\,d\theta$,

$$\left|\int_{C_2} \frac{dz}{z^2+1}\right| = \left|\int_0^\pi \frac{iR\exp(\theta i)}{1+R^2\exp(2\theta i)}\,d\theta\right| \le \int_0^\pi \frac{R}{R^2-1}\,d\theta, \tag{10.9.4}$$

which tends to zero as $R\to\infty$. On the other hand, the residue theorem gives

$$\oint_C \frac{dz}{z^2+1} = 2\pi i \operatorname{Res}\left(\frac{1}{z^2+1}; i\right) = 2\pi i \lim_{z\to i} \frac{z-i}{z^2+1} = 2\pi i \times \frac{1}{2i} = \pi. \tag{10.9.5}$$

Therefore,

$$\int_0^\infty \frac{dx}{x^2+1} = \frac{\pi}{2}. \tag{10.9.6}$$

Note that we only evaluated the residue in the upper half-plane because it is the only one inside the contour. □

This example illustrates the basic concepts of evaluating definite integrals by the residue theorem. We introduce a closed contour that includes the real axis and an additional contour. We must then evaluate the integral along this additional contour as well as the closed contour integral. If we properly choose our closed contour, this additional integral vanishes. For certain classes of general integrals, we shall now show that this additional contour is a circular arc at infinity.

Theorem: *If, on a circular arc C_R with a radius R and center at the origin, $zf(z) \to 0$ uniformly with $|z| \in C_R$ and as $R \to \infty$, then*

$$\lim_{R\to\infty} \int_{C_R} f(z)\,dz = 0. \tag{10.9.7}$$

The proof is as follows: If $|zf(z)| \le M_R$, then $|f(z)| \le M_R/R$. Because the length of C_R is αR, where α is the subtended angle,

$$\left| \int_{C_R} f(z)\,dz \right| \le \frac{M_R}{R}\,\alpha R = \alpha M_R \to 0, \tag{10.9.8}$$

because $M_R \to 0$ as $R \to \infty$. □

• **Example 10.9.2**

A simple illustration of this theorem is the integral

$$\int_{-\infty}^{\infty} \frac{dx}{x^2+x+1} = \int_{C_1} \frac{dz}{z^2+z+1}. \tag{10.9.9}$$

A quick check shows that $z/(z^2+z+1)$ tends to zero uniformly as $R \to \infty$. Therefore, if we use the contour pictured in Figure 10.9.1,

$$\int_{-\infty}^{\infty} \frac{dx}{x^2+x+1} = \oint_C \frac{dz}{z^2+z+1} = 2\pi i \operatorname{Res}\left(\frac{1}{z^2+z+1}; -\tfrac{1}{2} + \tfrac{\sqrt{3}}{2} i \right) \tag{10.9.10}$$

$$= 2\pi i \lim_{z \to -\frac{1}{2}+\frac{\sqrt{3}}{2}i} \left(\frac{1}{2z+1} \right) = \frac{2\pi}{\sqrt{3}}. \tag{10.9.11}$$

□

• **Example 10.9.3**

Let us evaluate

$$\int_0^{\infty} \frac{dx}{x^6+1}. \tag{10.9.12}$$

In place of an infinite semicircle in the upper half-plane, consider the following integral

$$\oint_C \frac{dz}{z^6+1}, \tag{10.9.13}$$

where we show the closed contour in Figure 10.9.2. We chose this contour for two reasons. First, we only have to evaluate one residue rather than the three enclosed in a traditional

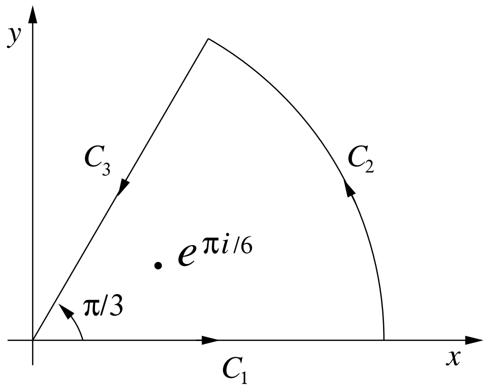


Figure 10.9.2: Contour used in evaluating the integral, Equation 10.9.13.

upper half-plane contour. Second, the contour integral along C_3 simplifies to a particularly simple and useful form.

Because the only enclosed singularity lies at $z = e^{\pi i/6}$,

$$\oint_C \frac{dz}{z^6+1} = 2\pi i \operatorname{Res}\left(\frac{1}{z^6+1}; e^{\pi i/6}\right) = 2\pi i \lim_{z\to e^{\pi i/6}} \frac{z-e^{\pi i/6}}{z^6+1} \tag{10.9.14}$$

$$= 2\pi i \lim_{z\to e^{\pi i/6}} \frac{1}{6z^5} = -\frac{\pi i}{3} e^{\pi i/6}. \tag{10.9.15}$$

Let us now evaluate Equation 10.9.12 along each of the legs of the contour:

$$\int_{C_1} \frac{dz}{z^6+1} = \int_0^\infty \frac{dx}{x^6+1}, \tag{10.9.16}$$

$$\int_{C_2} \frac{dz}{z^6+1} = 0, \tag{10.9.17}$$

because of Equation 10.9.7 and

$$\int_{C_3} \frac{dz}{z^6+1} = \int_\infty^0 \frac{e^{\pi i/3}\,dr}{r^6+1} = -e^{\pi i/3} \int_0^\infty \frac{dx}{x^6+1}, \tag{10.9.18}$$

since $z = re^{\pi i/3}$.

Substituting into Equation 10.9.15,

$$\left(1-e^{\pi i/3}\right) \int_0^\infty \frac{dx}{x^6+1} = -\frac{\pi i}{3} e^{\pi i/6} \tag{10.9.19}$$

or

$$\int_0^\infty \frac{dx}{x^6+1} = \frac{\pi i}{6} \frac{2ie^{\pi i/6}}{e^{\pi i/6}\left(e^{\pi i/6}-e^{-\pi i/6}\right)} = \frac{\pi}{6\sin(\pi/6)} = \frac{\pi}{3}. \tag{10.9.20}$$

□

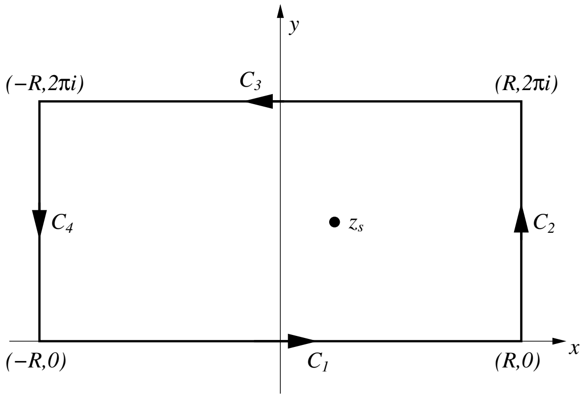


Figure 10.9.3: Rectangular closed contour used to obtain Equation 10.9.31.

• Example 10.9.4

Rectangular closed contours are best for the evaluation of integrals that involve hyperbolic sines and cosines. To illustrate[9] this, let us evaluate the integral

$$2\int_0^\infty \frac{\sin(ax)\,\sinh(x)}{[b+\cosh(x)]^2}\,dx = \int_{-\infty}^\infty \frac{\sin(ax)\,\sinh(x)}{[b+\cosh(x)]^2}\,dx = \Im\left[\int_{-\infty}^\infty \frac{\sinh(x)e^{iax}}{[b+\cosh(x)]^2}\,dx\right], \tag{10.9.21}$$

where $a > 0$ and $b > 1$.

We begin by determining the value of

$$\oint_C \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz$$

about the closed contour shown in Figure 10.9.3. Writing this contour integral in terms of the four line segments that constitute the closed contour, we have

$$\begin{aligned}\oint_C \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz &= \int_{C_1} \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz + \int_{C_2} \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz \\ &\quad + \int_{C_3} \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz + \int_{C_4} \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz.\end{aligned} \tag{10.9.22}$$

Because the integrand behaves as e^{-R} as $R \to \infty$, the integrals along C_2 and C_4 vanish. On the other hand,

$$\int_{C_1} \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz = \int_{-\infty}^\infty \frac{\sinh(x)e^{iax}}{[b+\cosh(x)]^2}\,dx, \tag{10.9.23}$$

and

$$\int_{C_3} \frac{\sinh(z)e^{iaz}}{[b+\cosh(z)]^2}\,dz = -e^{-2\pi a}\int_{-\infty}^\infty \frac{\sinh(x)e^{iax}}{[b+\cosh(x)]^2}\,dx, \tag{10.9.24}$$

because $\cosh(x+2\pi i) = \cosh(x)$ and $\sinh(x+2\pi i) = \sinh(x)$.

---

[9] This is a slight variation on a problem solved by Spyrou, K. J., B. Cotton, and B. Gurd, 2002: Analytical expressions of capsize boundary for a ship with roll bias in beam waves. *J. Ship Res.*, **46**, 167–174.

Within the closed contour C, we have a single singularity where $b + \cosh(z_s) = 0$ or $e^{z_s} = -b - \sqrt{b^2 - 1}$ or $z_s = \ln(b + \sqrt{b^2 - 1}\,) + \pi i$. To discover the nature of this singularity, we expand $b + \cosh(z)$ in a Taylor expansion and find that

$$b + \cosh(z) = \sinh(z_s)(z - z_s) + \tfrac{1}{2}\cosh(z_s)(z - z_s)^2 + \cdots. \tag{10.9.25}$$

Therefore, we have a second-order pole at $z = z_s$. Therefore, the value of the residue there is

$$\operatorname{Res}\left[\frac{\sinh(z)e^{iaz}}{[b + \cosh(z)]^2}; z_s\right] = \lim_{z \to z_s} \frac{d}{dz}\left[\frac{\sinh(z)e^{iaz}}{\sinh^2(z_s) + \sinh(z_s)\cosh(z_s)(z - z_s) + \cdots}\right] \tag{10.9.26}$$

$$= \frac{ia\,e^{-\pi a}}{\sinh(z_s)} \exp[ia\cosh^{-1}(b)]. \tag{10.9.27}$$

Therefore,

$$\int_{-\infty}^{\infty} \frac{\sinh(x)e^{iax}}{[b + \cosh(x)]^2}\,dx = -\frac{2\pi a \exp[-\pi a + ai\cosh^{-1}(b)]}{(1 - e^{-2\pi a})\sinh(z_s)} = \frac{\pi a \exp[ai\cosh^{-1}(b)]}{\sqrt{b^2 - 1}\,\sinh(\pi a)}, \tag{10.9.28}$$

because

$$\sinh(z_s) = \frac{1}{2}\left[-b - \sqrt{b^2 - 1} + \frac{1}{b + \sqrt{b^2 - 1}}\right] = -\sqrt{b^2 - 1}. \tag{10.9.29}$$

Substituting Equation 10.9.28 into Equation 10.9.21 yields

$$\int_0^{\infty} \frac{\sin(ax)\,\sinh(x)}{[b + \cosh(x)]^2}\,dx = \frac{\pi a \sin[a\cosh^{-1}(b)]}{2\sqrt{b^2 - 1}\,\sinh(\pi a)}. \tag{10.9.30}$$

□

• Example 10.9.5

The method of residues is also useful in the evaluation of definite integrals of the form $\int_0^{2\pi} F[\sin(\theta), \cos(\theta)]\,d\theta$, where F is a quotient of polynomials in $\sin(\theta)$ and $\cos(\theta)$. For example, let us evaluate the integral[10]

$$I = \int_0^{2\pi} \frac{\cos^3(\theta)}{\cos^2(\theta) - a^2}\,d\theta, \qquad a > 1. \tag{10.9.31}$$

We begin by introducing the complex variable $z = e^{i\theta}$. This substitution yields the closed contour integral

$$I = \frac{1}{2i} \oint_C \frac{(z^2 + 1)^3}{(z^2 + 1)^2 - 4a^2z^2}\,\frac{dz}{z^2}, \tag{10.9.32}$$

where C is a circle of radius 1 taken in the positive sense. The integrand of Equation 10.9.32 has five singularities: a second-order pole at $z_5 = 0$ and simple poles located at

$$z_1 = -a - \sqrt{a^2 - 1}, \qquad z_2 = -a + \sqrt{a^2 - 1}, \tag{10.9.33}$$

[10] Simplified version of an integral presented by Jiang, Q. F., and R. B. Smith, 2000: V-waves, bow shocks, and wakes in supercritical hydrostatic flow. *J. Fluid Mech.*, **406**, 27–53.

$$z_3 = a - \sqrt{a^2-1}, \quad \text{and} \quad z_4 = a + \sqrt{a^2-1}. \tag{10.9.34}$$

Only the singularities z_2, z_3, and z_5 lie within C. Consequently, the value of I equals $2\pi i$ times the sum of the residues at these three singularities. The residues equal

$$\begin{aligned}
&\operatorname{Res}\left\{\frac{(z^2+1)^3}{z^2[(z^2+1)^2-4a^2z^2]}; -a+\sqrt{a^2-1}\right\} \\
&= \lim_{z\to -a+\sqrt{a^2-1}} \frac{(z^2+1)^3}{z^2} \lim_{z\to -a+\sqrt{a^2-1}} \frac{z+a-\sqrt{a^2-1}}{(z^2+1)^2-4a^2z^2} && (10.9.35)\\
&= \lim_{z\to -a+\sqrt{a^2-1}} \frac{(z^2+1)^3}{4z^3(z^2+1-2a^2)} && (10.9.36)\\
&= -\frac{a^2(a-\sqrt{a^2-1}\,)^3}{(2a^2-1-2a\sqrt{a^2-1}\,)(a^2-1-a\sqrt{a^2-1}\,)}, && (10.9.37)
\end{aligned}$$

$$\begin{aligned}
&\operatorname{Res}\left\{\frac{(z^2+1)^3}{z^2[(z^2+1)^2-4a^2z^2]}; a-\sqrt{a^2-1}\right\} \\
&= \lim_{z\to a-\sqrt{a^2-1}} \frac{(z^2+1)^3}{z^2} \lim_{z\to a-\sqrt{a^2-1}} \frac{z-a+\sqrt{a^2-1}}{(z^2+1)^2-4a^2z^2} && (10.9.38)\\
&= \lim_{z\to a-\sqrt{a^2-1}} \frac{(z^2+1)^3}{4z^3(z^2+1-2a^2)} && (10.9.39)\\
&= \frac{a^2(a-\sqrt{a^2-1}\,)^3}{(2a^2-1-2a\sqrt{a^2-1}\,)(a^2-1-a\sqrt{a^2-1}\,)}, && (10.9.40)
\end{aligned}$$

and

$$\begin{aligned}
&\operatorname{Res}\left\{\frac{(z^2+1)^3}{z^2[(z^2+1)^2-4a^2z^2]}; 0\right\} \\
&= \lim_{z\to 0} \frac{d}{dz}\left[\frac{(z^2+1)^3}{(z^2+1)^2-4a^2z^2}\right] && (10.9.41)\\
&= \lim_{z\to 0} \frac{6z[(z^2+1)^4-4a^2z^2(z^2+1)^2]-4z(z^2+1)^3(z^2+1-2a^2)}{[(z^2+1)^2-4a^2z^2]^2} && (10.9.42)\\
&= 0. && (10.9.43)
\end{aligned}$$

Summing the residues, we obtain 0. Therefore,

$$\int_0^{2\pi} \frac{\cos^3(\theta)}{\cos^2(\theta)-a^2}\,d\theta = 0, \qquad a>1. \tag{10.9.44}$$

Problems

Use the residue theorem to verify the following integrals:

1. $\displaystyle\int_0^\infty \frac{dx}{x^4+1} = \frac{\pi\sqrt{2}}{4}$

2. $\displaystyle\int_{-\infty}^\infty \frac{dx}{(x^2+4x+5)^2} = \frac{\pi}{2}$

3. $\displaystyle\int_{-\infty}^{\infty} \frac{x\,dx}{(x^2+1)(x^2+2x+2)} = -\frac{\pi}{5}$

4. $\displaystyle\int_0^{\infty} \frac{x^2}{x^6+1}\,dx = \frac{\pi}{6}$

5. $\displaystyle\int_0^{\infty} \frac{dx}{(x^2+1)^2} = \frac{\pi}{4}$

6. $\displaystyle\int_0^{\infty} \frac{dx}{(x^2+1)(x^2+4)^2} = \frac{5\pi}{288}$

7.
$$\int_{-\infty}^{\infty} \frac{x^2\;dx}{(x^2+a^2)(x^2+b^2)^2} = \frac{\pi}{2b(a+b)^2}, \qquad a, b > 0$$

8.
$$\int_0^{\infty} \frac{t^2}{(t^2+1)[t^2(a/h+1)+(a/h-1)]}\,dt = \frac{\pi}{4}\left[1-\sqrt{\frac{a-h}{a+h}}\right], \quad a > h$$

9.
$$\int_0^{\pi/2} \frac{d\theta}{a+\sin^2(\theta)} = \frac{\pi}{2\sqrt{a+a^2}}, \qquad a > 0$$

10.
$$\int_0^{\pi/2} \frac{d\theta}{a^2\cos^2(\theta)+b^2\sin^2(\theta)} = \frac{\pi}{2ab}, \qquad b \geq a > 0$$

11.
$$\int_0^{\pi} \frac{\sin^2(\theta)}{a+b\cos(\theta)}\,d\theta = \frac{\pi}{b^2}\left(a-\sqrt{a^2-b^2}\right), \qquad a > b > 0$$

12.
$$\int_0^{2\pi} \frac{e^{in\theta}}{1+2r\cos(\theta)+r^2}\,d\theta = 2\pi\frac{(-r)^n}{1-r^2}, \qquad 1 > |r|, \quad n = 0, 1, 2, \ldots$$

13.
$$\int_0^{2\pi} \sin^{2n}(\theta)\,d\theta = \frac{2\pi(2n)!}{(2^n n!)^2}$$

14.
$$\int_{-\pi}^{\pi} \frac{\cos(n\theta)}{\cos(\theta)+\alpha}\,d\theta = 2\pi\frac{(-\alpha+\sqrt{\alpha^2-1}\,)^n}{\sqrt{\alpha^2-1}}, \quad \alpha > 1, \quad n \geq 0$$

Hint:
$$\frac{2\sqrt{\alpha^2-1}}{z^2+2\alpha z+1} = \frac{1}{z+\alpha-\sqrt{\alpha^2-1}} - \frac{1}{z+\alpha+\sqrt{\alpha^2-1}}$$

15.
$$\int_0^{\pi} \frac{\cos(n\theta)}{\cosh(\alpha)-\cos(\theta)}\,d\theta = \frac{\pi}{\sinh(\alpha)}e^{-n\alpha}, \quad \alpha \neq 0, \quad n \geq 0$$

Hint: See Example 10.7.5.

16. Show that
$$\int_0^{\infty} \frac{x^2}{(1-x^2)^2+a^2x^2}\,dx = \frac{\pi}{2|a|},$$
where a is real and not equal to zero. Hint: Show that the poles of
$$f(z) = \frac{z^2}{(1-z^2)^2+a^2z^2}$$

are simple and equal

$$z_n = \begin{cases} \pm\frac{1}{2}\left(\pm\sqrt{4-a^2}+|a|i\right), & \text{if} \quad 0<|a|<2, \\ \pm\frac{i}{2}\left(|a|\pm\sqrt{a^2-4}\right), & \text{if} \quad 2<|a|. \end{cases}$$

If $|a| = 2$, we have second-order poles at $z_n = \pm i$.

17. Show that

$$\int_0^\infty \frac{\cos(ax)}{\cosh^2(bx)}\,dx = \frac{\pi a}{2b^2 \sinh[a\pi/(2b)]}, \qquad a, b > 0.$$

Hint: Evaluate the closed contour integral

$$\oint_C \frac{e^{iaz}}{\cosh^2(bz)}\,dz,$$

where C is a *rectangular* contour with vertices at $(\infty, 0)$, $(-\infty, 0)$, $(\infty, \pi/b)$, and $(-\infty, \pi/b)$.

18. Show[11] that

$$\int_0^\infty \frac{dx}{\cosh(x)\cosh(x+a)} = \begin{cases} 2a/\sinh(a), & \text{if} \quad a \neq 0, \\ 2, & \text{if} \quad a = 2. \end{cases}$$

Hint: Evaluate the closed contour integral

$$\oint_C \frac{z}{\cosh(z)\cosh(z+a)}\,dz,$$

where C is a *rectangular* contour with vertices at $(\infty, 0)$, $(-\infty, 0)$, (∞, π), and $(-\infty, \pi)$.

19. During an electromagnetic calculation, Strutt[12] needed to prove that

$$\pi\frac{\sinh(\sigma x)}{\cosh(\sigma\pi)} = 2\sigma \sum_{n=0}^{\infty} \frac{\cos\left[\left(n+\frac{1}{2}\right)(x-\pi)\right]}{\sigma^2+\left(n+\frac{1}{2}\right)^2}, \qquad |x| \le \pi.$$

Verify his proof by doing the following:

Step 1: Using the residue theorem, show that

$$\frac{1}{2\pi i}\oint_{C_N} \pi\frac{\sinh(xz)}{\cosh(\pi z)}\frac{dz}{z-\sigma} = \pi\frac{\sinh(\sigma x)}{\cosh(\sigma\pi)} - \sum_{n=-N-1}^{N} \frac{(-1)^n \sin\left[\left(n+\frac{1}{2}\right)x\right]}{\sigma - i\left(n+\frac{1}{2}\right)},$$

where C_N is a circular contour that includes the poles $z = \sigma$ and $z_n = \pm i\left(n+\frac{1}{2}\right)$, $n = 0, 1, 2, \ldots, N$.

[11] See Yan, J. R., X. H. Yan, J. Q. You, and J. X. Zhong, 1993: On the interaction between two nonpropagating hydrodynamic solitons. *Phys. Fluids A*, **5**, 1651–1656.

[12] Strutt, M. J. O., 1934: Berechnung des hochfrequenten Feldes einer Kreiszylinderspule in einer konzentrischen leitenden Schirmhülle mit ebenen Deckeln. *Hochfrequenztechn. Elecktroak.*, **43**, 121–123.

Step 2: Show that in the limit of $N \to \infty$, the contour integral vanishes. Hint: Examine the behavior of $z\ \sinh(xz)/[(z-\sigma)\cosh(\pi z)]$ as $|z| \to \infty$. Use Equation 10.9.7 where C_R is the circular contour.

Step 3: Break the infinite series in Step 1 into two parts and simplify.

In the chapter on Fourier series, we shall show how we can obtain the same series by direct integration.

10.10 CAUCHY'S PRINCIPAL VALUE INTEGRAL

The conventional definition of the integral of a function $f(x)$ of the real variable x over a finite interval $a \le x \le b$ assumes that $f(x)$ has a definite finite value at each point within the interval. We shall now extend this definition to cover cases when $f(x)$ is infinite at a finite number of points within the interval.

Consider the case when there is only one point c at which $f(x)$ becomes infinite. If c is not an endpoint of the interval, we take two small positive numbers ϵ and η and examine the expression

$$\int_a^{c-\epsilon} f(x)\,dx + \int_{c+\eta}^b f(x)\,dx. \tag{10.10.1}$$

If Equation 10.10.1 exists and tends to a unique limit as ϵ and η tend to zero independently, we say that the improper integral of $f(x)$ over the interval exists, its value being defined by

$$\int_a^b f(x)\,dx = \lim_{\epsilon\to 0}\int_a^{c-\epsilon} f(x)\,dx + \lim_{\eta\to 0}\int_{c+\eta}^b f(x)\,dx. \tag{10.10.2}$$

If, however, the expression does not tend to a limit as ϵ and η tend to zero independently, it may still happen that

$$\lim_{\epsilon\to 0}\left\{\int_a^{c-\epsilon} f(x)\,dx + \int_{c+\epsilon}^b f(x)\,dx\right\} \tag{10.10.3}$$

exists. When this is the case, we call this limit the *Cauchy principal value* of the improper integral and denote it by

$$PV\int_a^b f(x)\,dx. \tag{10.10.4}$$

Finally, if $f(x)$ becomes infinite at an endpoint, say a, of the range of integration, we say that $f(x)$ is integrable over $a \le x \le b$ if

$$\lim_{\epsilon\to 0^+}\int_{a+\epsilon}^b f(x)\,dx \tag{10.10.5}$$

exists.

• Example 10.10.1

Consider the integral $\int_{-1}^2 dx/x$. This integral does not exist in the ordinary sense because of the strong singularity at the origin. However, the integral would exist if

$$\lim_{\epsilon\to 0}\int_{-1}^{\epsilon}\frac{dx}{x} + \lim_{\delta\to 0}\int_{\delta}^2\frac{dx}{x} \tag{10.10.6}$$

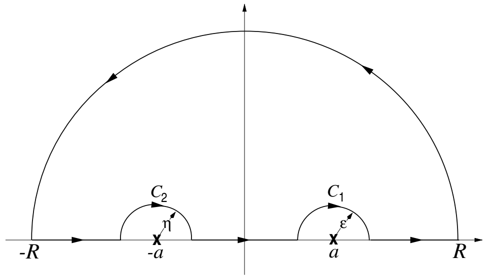


Figure 10.10.1: Contour C used in Example 10.10.2.

existed and had a unique value as ϵ and δ independently approach zero. Because this limit equals

$$\lim_{\epsilon,\delta\to 0}\left[\ln(\epsilon)+\ln(2)-\ln(\delta)\right]=\lim_{\epsilon,\delta\to 0}\left[\ln(2)-\ln(\delta/\epsilon)\right], \tag{10.10.7}$$

our integral would have the value of $\ln(2)$ if $\delta=\epsilon$. This particular limit is the Cauchy principal value of the improper integral, which we express as

$$PV\int_{-1}^{2}\frac{dx}{x}=\ln(2). \tag{10.10.8}$$

□

We can extend these ideas to complex integrals used to determine the value or principal value of an improper integral by Cauchy's residue theorem when the integrand has a singularity on the contour of integration. We avoid this difficulty by deleting from the area within the contour that portion which also lies within a small circle $|z-c|=\epsilon$ and then integrate around the boundary of the remaining region. This process is called *indenting* the contour.

The integral around the indented contour is calculated by the theorem of residues and then the radius of each indentation is made to tend to zero. This process gives the Cauchy principal value of the improper integral. The details of this method are shown in the following examples.

• Example 10.10.2

Let us show that

$$PV\int_{-\infty}^{\infty}\frac{\cos(x)}{a^2-x^2}\,dx=\frac{\pi\,\sin(a)}{a},\qquad a>0. \tag{10.10.9}$$

Consider the integral

$$\oint_C \frac{e^{iz}}{a^2-z^2}\,dz, \tag{10.10.10}$$

where the closed contour C consists of the real axis from $-R$ to R and a semicircle in the upper half of the z-plane where this segment is its diameter. See Figure 10.10.1. Because the integrand has poles at $z=\pm a$, which lie on this contour, we modify C by making an indentation of radius ϵ at a and another of radius η at $-a$. The integrand is now analytic within and on C and Equation 10.10.10 equals zero by the Cauchy-Goursat theorem.

Evaluating each part of the integral, Equation 10.10.10, we have that

$$\int_0^\pi \frac{e^{iR\cos(\theta)-R\sin(\theta)}}{a^2-R^2e^{2\theta i}} iRe^{\theta i}\,d\theta + \int_{C_1} \frac{e^{iz}}{a^2-z^2}\,dz + \int_{C_2} \frac{e^{iz}}{a^2-z^2}\,dz + \int_{-R}^{-a-\eta} \frac{e^{ix}}{a^2-x^2}\,dx + \int_{-a+\eta}^{a-\epsilon} \frac{e^{ix}}{a^2-x^2}\,dx + \int_{a-\epsilon}^{R} \frac{e^{ix}}{a^2-x^2}\,dx = 0, \qquad \textbf{(10.10.11)}$$

where C_1 and C_2 denote the integrals around the indentations at a and $-a$, respectively. The modulus of the first term on the left side of Equation 10.10.11 is less than $\pi R/(R^2-a^2)$ so that this term tends to zero as $R\to\infty$. To evaluate C_1, we observe that $z = a + \epsilon e^{\theta i}$ along C_1, where θ decreases from π to 0. Hence,

$$\int_{C_1} \frac{e^{iz}}{a^2-z^2}\,dz = \lim_{\epsilon\to 0}\int_\pi^0 \exp\left(ia+i\epsilon e^{\theta i}\right) \frac{\epsilon i e^{\theta i}}{-2a\epsilon e^{\theta i}-\epsilon^2 e^{2\theta i}}\,d\theta \qquad \textbf{(10.10.12)}$$

$$= \lim_{\epsilon\to 0}\int_0^\pi \exp\left(ia+i\epsilon e^{\theta i}\right) \frac{i}{2a+\epsilon e^{\theta i}}\,d\theta = \frac{\pi i e^{ia}}{2a}. \qquad \textbf{(10.10.13)}$$

Similarly,

$$\int_{C_2} \frac{e^{iz}}{a^2-z^2}\,dz = -\frac{\pi i e^{-ia}}{2a}, \qquad \textbf{(10.10.14)}$$

as η tends to zero.

Upon letting $R\to\infty$, $\epsilon\to 0$, and $\eta\to 0$, we find that

$$PV\int_{-\infty}^{\infty} \frac{e^{ix}}{a^2-x^2}\,dx = -\frac{\pi i}{2a}\left(e^{ia}-e^{-ia}\right) = \frac{\pi\,\sin(a)}{a}. \qquad \textbf{(10.10.15)}$$

Finally, equating the real and imaginary parts, we obtain

$$PV\int_{-\infty}^{\infty} \frac{\cos(x)}{a^2-x^2}\,dx = \frac{\pi\,\sin(a)}{a}, \quad PV\int_{-\infty}^{\infty} \frac{\sin(x)}{a^2-x^2}\,dx = 0. \qquad \textbf{(10.10.16)}$$

□

- **Example 10.10.3**

Let us show that

$$\int_{-\infty}^{\infty} \frac{\sin(x)}{x}\,dx = \pi. \qquad \textbf{(10.10.17)}$$

Consider the integral

$$\oint_C \frac{e^{iz}}{z}\,dz, \qquad \textbf{(10.10.18)}$$

where the closed contour C consists of the real axis from $-R$ to R and a semicircle in the upper half of the z-plane where this segment is its diameter. Because the integrand has a pole at $z=0$, which lies on the contour, we modify C by making an indentation of radius ϵ at $z=0$. See Figure 10.10.2. Because e^{iz}/z is analytic along C,

$$\int_0^\pi e^{iR\cos(\theta)-R\sin(\theta)} i\,d\theta + \int_{-R}^{-\epsilon} \frac{e^{ix}}{x}\,dx + \int_{C_1} \frac{e^{iz}}{z}\,dz + \int_\epsilon^R \frac{e^{ix}}{x}\,dx = 0. \qquad \textbf{(10.10.19)}$$

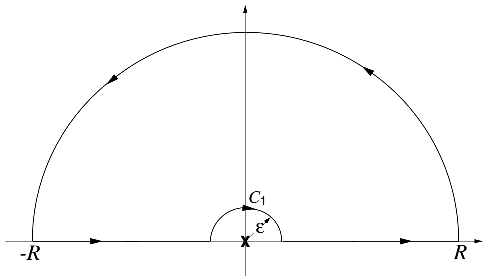


Figure 10.10.2: Contour C used in Example 10.10.3.

Since $e^{-R\sin(\theta)} < e^{-R\theta}$ for $0 < \theta < \pi$,

$$\left|\int_0^\pi e^{iR\cos(\theta)-R\sin(\theta)} i\, d\theta\right| \le \int_0^\pi e^{-R\theta}\, d\theta = \frac{1-e^{-\pi R}}{R}, \tag{10.10.20}$$

which tends to zero as $R \to \infty$. Therefore,

$$\int_{-\infty}^{-\epsilon} \frac{e^{ix}}{x}\, dx + \int_\epsilon^\infty \frac{e^{ix}}{x}\, dx = -\int_{C_1} \frac{e^{iz}}{z}\, dz. \tag{10.10.21}$$

Now,

$$\int_{C_1} \frac{e^{iz}}{z}\, dz = \int_{C_1} \frac{dz}{z} + i\int_{C_1} dz - \int_{C_1} \frac{z}{2}\, dz + \cdots = -\pi i \tag{10.10.22}$$

in the limit $\epsilon \to 0$ because $z = \epsilon e^{\theta i}$. Consequently, in the limit of $\epsilon \to 0$,

$$PV \int_{-\infty}^\infty \frac{e^{ix}}{x}\, dx = \pi. \tag{10.10.23}$$

Upon separating the real and imaginary parts, we obtain

$$PV \int_{-\infty}^\infty \frac{\cos(x)}{x}\, dx = 0, \qquad \int_{-\infty}^\infty \frac{\sin(x)}{x}\, dx = \pi. \tag{10.10.24}$$

Problems

1. Noting that

$$\int_0^{\theta-\epsilon} \frac{d\varphi}{\cos(\varphi)-\cos(\theta)} = \frac{1}{\sin(\theta)} \ln\left|\frac{\sin\left[\frac{1}{2}(\theta+\varphi)\right]}{\sin\left[\frac{1}{2}(\theta-\varphi)\right]}\right|_0^{\theta-\epsilon},$$

and

$$\int_{\theta+\epsilon}^\pi \frac{d\varphi}{\cos(\varphi)-\cos(\theta)} = \frac{1}{\sin(\theta)} \ln\left|\frac{\sin\left[\frac{1}{2}(\theta+\varphi)\right]}{\sin\left[\frac{1}{2}(\theta-\varphi)\right]}\right|_{\theta+\epsilon}^{\pi},$$

show that

$$PV \int_0^\pi \frac{d\varphi}{\cos(\varphi)-\cos(\theta)} = 0, \qquad 0 < \theta < \pi.$$

2. Using $f(z) = e^{i\pi z/2}/(z^2-1)$, show that

$$\int_{-\infty}^{\infty} \frac{\cos(\pi x/2)}{x^2-1}\,dx = -\pi.$$

3. Show that

$$\int_{-\infty}^{\infty} \frac{e^{ax}-e^{bx}}{1-e^x}\,dx = \pi[\cot(a\pi)-\cot(b\pi)], \qquad 0 < a, b < 1.$$

Use a rectangular contour with vertices at $(-R,0)$, $(R,0)$, $(-R,\pi)$, and (R,π) with a semicircle indentation at the origin.

4. Show[13] that

$$\int_{-\infty}^{\infty} \frac{1-\cos[2a(x+\zeta)]}{(x+\zeta)^2(x^2+\alpha^2)}\,dx = \frac{\pi}{\alpha(\zeta^2+\alpha^2)^2}\big\{2a\alpha(\zeta^2+\alpha^2)+(\zeta^2-\alpha^2)$$
$$- e^{-2a\alpha}\left[(\zeta^2-\alpha^2)\cos(2a\zeta)+2\alpha\zeta\sin(2a\zeta)\right]\big\},$$

where a, α, and ζ are real. Use a semicircular contour of infinite radius with the real axis as its diameter.

5. Using the complex function $e^{imz}/(z-a)$ and a closed contour similar to that shown in Figure 10.10.2, show that

$$PV\int_{-\infty}^{\infty} \frac{\cos(mx)}{x-a}\,dx = -\pi\sin(ma), \qquad \text{and} \qquad PV\int_{-\infty}^{\infty} \frac{\sin(mx)}{x-a}\,dx = \pi\cos(ma),$$

where $m > 0$ and a is real.

6. Using a closed contour similar to that shown in Figure 10.10.2, except that we now have two small semicircles around the singularities on the real axis, show that

$$PV\int_{-\infty}^{\infty} \frac{xe^{xi}}{x^2-\pi^2}\,dx = -\pi i, \qquad \text{and} \qquad PV\int_{-\infty}^{\infty} \frac{e^{imx}}{(x-1)(x-3)}\,dx = \frac{\pi i}{2}\left(e^{3mi}-e^{mi}\right),$$

where $m > 0$.

7. Redo Example 10.10.3 except the contour is now a rectangle with vertices at $\pm R$ and $\pm R + Ri$ indented at the origin.

8. Let us show[14] that

$$G(\alpha) = PV\int_{-1}^{1} \frac{dx}{(x+\alpha)\sqrt{1-x^2}} = \begin{cases} \dfrac{\alpha\pi}{|\alpha|\sqrt{\alpha^2-1}}, & |\alpha| > 1, \\ 0, & |\alpha| < 1. \end{cases}$$

[13] Ko, S. H., and A. H. Nuttall, 1991: Analytical evaluation of flush-mounted hydrophone array response to the Corcos turbulent wall pressure spectrum. *J. Acoust. Soc. Am.*, **90**, 579–588.

[14] Ott, E., T. M. Antonsen, and R. V. Lovelace, 1977: Theory of foil-less diode generation of intense relativistic electron beams. *Phys. Fluids*, **20**, 1180–1184.

Step 1: Using the transformation $2ix = z - z^{-1}$, show that

$$G(\alpha) = PV \oint_{|z|=1} \frac{dz}{z^2 + 2i\alpha z - 1},$$

which has singularities at $z = \pm\sqrt{1-\alpha^2} + \alpha i$.

Step 2: To evaluate $G(\alpha)$ given in Step 1, the principal value can be evaluated using

$$G(\alpha) = \oint_{|z|=1-\epsilon} \frac{dz}{z^2 + 2i\alpha z - 1} + \oint_{|z|=1+\epsilon} \frac{dz}{z^2 + 2i\alpha z - 1},$$

where $\epsilon \to 0^+$. Use the residue theorem and evaluate these contour integrals, yielding the desired result.

9. Let the function $f(z)$ possess a simple pole with a residue $\text{Res}[f(z); c]$ on a simply closed contour C. If C is indented at c, show that the integral of $f(z)$ around the indentation tends to $-\text{Res}[f(z); c]\alpha i$ as the radius of the indentation tends to zero, α being the internal angle between the two parts of C meeting at c.

10.11 CONFORMAL MAPPING

Conformal mapping is a powerful technique for finding solutions, or for simplifying the process of finding solutions, to Laplace's differential equation in two dimensions. This method involves introducing two complex variables: $z = x + iy$ and $\tau = \rho + i\sigma$. These two complex variables are related to each other via the mapping $z = f(\tau)$. Under this mapping the Argand diagram for the z-variable is mapped into one for the τ-variable. In certain cases, for example $\tau = \sqrt{z}$, the complex z-plane may only map into a portion of the τ-plane. In other cases, say $\tau = z + 3i$, the complete z-plane would be mapped into the complete τ-plane.

Once we map the original domain into a simpler geometry (a half-plane, circle or square), how do we find the solution? There are several techniques available. One method, for example, recalls that the real and imaginary part of an analytic function satisfies Laplace's equation. Therefore, if we could construct an analytic function whose real or imaginary part satisfies the boundary conditions in the new domain, we would have the solution in the τ-plane. Then we could use the transformation to obtain the solution in the original z-plane.

What types of functions $f(z)$ are useful? Consider an arbitrary point z_0 in the complex z-plane. Assuming that $f'(z_0) \neq 0$, a straightforward transformation yields

$$\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)_{z_0} = |f'(z_0)|^2 \left(\frac{\partial^2 v}{\partial \rho^2} + \frac{\partial^2 v}{\partial \sigma^2}\right)_{\tau_0}, \qquad (\mathbf{10.11.1})$$

where $u(x, y)$ and $v(\rho, \sigma)$ are solutions to Laplace's equation in the z and τ planes, respectively. Thus, $f(z)$ must be analytic.

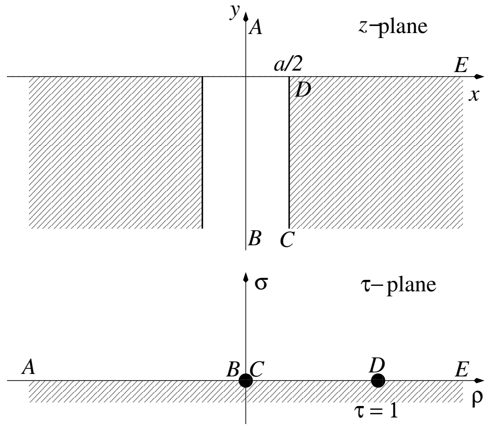


Figure 10.11.1: The conformal mapping used to find the fields of a semi-infinite ring head with a finite gap of width a. The potential on the right pole face equals 1 while the potential of the left pole face equals -1. In the z-plane the point A is located at $(0, \infty)$ while point B is located at $(0, -\infty)$. Because of symmetry the potential along the center of the gap AB equals 0.

• Example 10.11.1

In their study of magnetic recording, Curland and Judy[15] modeled the ring heads as two semi-infinite regions located below the x-axis and running to the right of $x = a/2$ and to the left of $x = -a/2$. See Figure 10.11.1.

From symmetry we need only consider the half-space $x > 0$. Consequently, the new boundary consists of the four line segments: AB, BC, CD and DE. If we require that the point D in the τ-plane lies at $\tau = 1$, we shall show in Example 10.11.7 that the desired conformal mapping is

$$z = \frac{a}{\pi}\left[\sqrt{\tau - 1} - \frac{i}{2}\log\left(\frac{1 - i\sqrt{\tau - 1}}{1 + i\sqrt{\tau - 1}}\right)\right] + \frac{a}{2}. \tag{10.11.2}$$

A useful method for illustrating this conformal mapping is to draw lines of constant ρ and σ in the z-plane. See Figure 10.11.2. This figure shows the local orthogonality between lines of constant ρ and σ.

The greatest difficulty in creating this figure was computing τ for a given z. This was done using the Newton-Raphson method. Starting at the top of the domain, the first guess there was given by $\tau = 1 + \pi^2 z^2$. Marching downward, the τ from the previous grid point was used for the initial guess. The corresponding MATLAB script is:

```
clear; delta = 0.01; % resolution of the grid

for jj = 1:201
for ii = 1:201
  XX(jj,ii) = delta*ii; YY(jj,ii) = delta*(jj-101);
  RHO(jj,ii) = NaN; SIGMA(jj,ii) = NaN;
end; end

% code for the domain x,y > 0
```

[15] Curland, N., and J. H. Judy, 1986: Calculation of exact ring head fields using conformal mapping. *IEEE Trans. Magnet.*, **MAG-22**, 1901–1903.

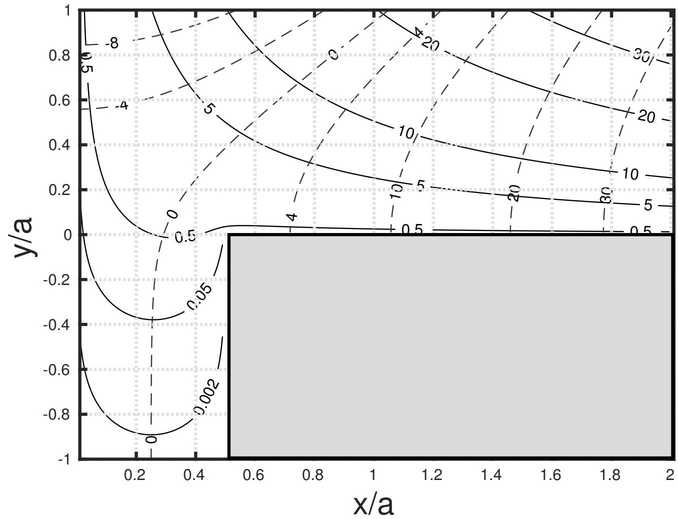


Figure 10.11.2: Lines of constant ρ (dashed lines) and σ (solid lines) given by the conformal mapping expressed by Equation 10.11.2.

```
for jj = 1:100
  y = 1 - delta*(jj-1);
for ii = 1:201
  x = delta*ii; z = complex(x,y);
  if (jj == 1) tau = 1+pi*pi*z*z; else tau = TAU(ii); end
  for icount = 1:10
    temp1 = sqrt(tau-1);
    temp2 = temp1 - 0.5*i*log(1-i*temp1) + 0.5*i*log(1+i*temp1);
    ff = temp2/pi + 0.5 - z; deriv = temp1 /(2*pi*tau);
    temp3 = ff/deriv; tau = tau - temp3; % Newton-Raphson method
  end
  TAU(ii) = tau; RHO(202-jj,ii) = real(tau);
  SIGMA(202-jj,ii) = imag(tau);
end; end
```

% code for the domain $0 < x < \frac{1}{2}$ and $y < 0$

```
for jj = 1:101
  y = delta - delta*jj;
for ii = 1:49
  x = delta*ii; z = complex(x,y);
  tau = TAU(ii); % first guess
  for icount = 1:10
    temp1 = sqrt(tau-1);
    temp2 = temp1 - 0.5*i*log(1-i*temp1) + 0.5*i*log(1+i*temp1);
    ff = temp2/pi + 0.5 - z; deriv = temp1 /(2*pi*tau);
    temp3 = ff/deriv; tau = tau - temp3; % Newton-Raphson method
  end
```

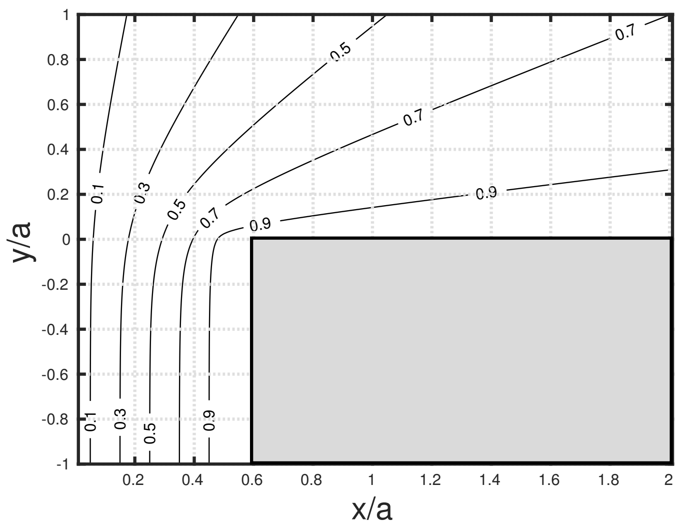


Figure 10.11.3: The solution to Laplace's equation when the left boundary is held at 0 while the left and top sides of the shaded rectangle are held at 1. This figure shows only a portion of the domain $x > 0$ and $|y| < \infty$.

```
  TAU(ii) = tau; RHO(102-jj,ii) = real(tau);
  SIGMA(102-jj,ii) = imag(tau);
end; end

% plot the conformal mapping Equation 10.11.2

figure
[C,h] = contour(XX,YY,SIGMA,[0.002,0.05,0.5,5,10,20,30],'k');
clabel(C,h,'FontSize',10,'Color','k','Rotation',0)
xlabel('x/a','FontSize',20); ylabel('y/a','FontSize',20);
hold on
v = [-8,-4,0,4,10,20,30];
[C,h] = contour(XX,YY,RHO,v,'--b');
clabel(C,h,'FontSize',10,'Color','b','Rotation',0)
```

Now that we can transform between the z-plane and the τ-plane, and vice versa, let us turn our attention to finding the solution to Laplace's equation in the τ-plane. There the solution equals 1 for $\rho > 0$ and 0 for $\rho < 0$ along $\sigma = 0$.

Consider now the analytic function (except at the branch point $\tau = 0$)

$$f(\tau) = i - \log(\tau)/\pi. \qquad \textbf{(10.11.3)}$$

A quick check (using $\tau = re^{i\theta}$) shows that the *imaginary* part of $f(\tau)$, $v(r,\theta) = 1 - \theta/\pi$, satisfies Laplace's equation and the boundary conditions. Thus, constructing the solution is as follows: For a given x and y, we use our MATLAB code to compute τ. Substituting that τ into Equation 10.11.3 we compute $f(\tau)$. Taking the imaginary part, we have the solution at x and y. Figure 10.11.3 illustrates the solution for the domain $0 < x < 2$ and $-1 < y < 1$. □

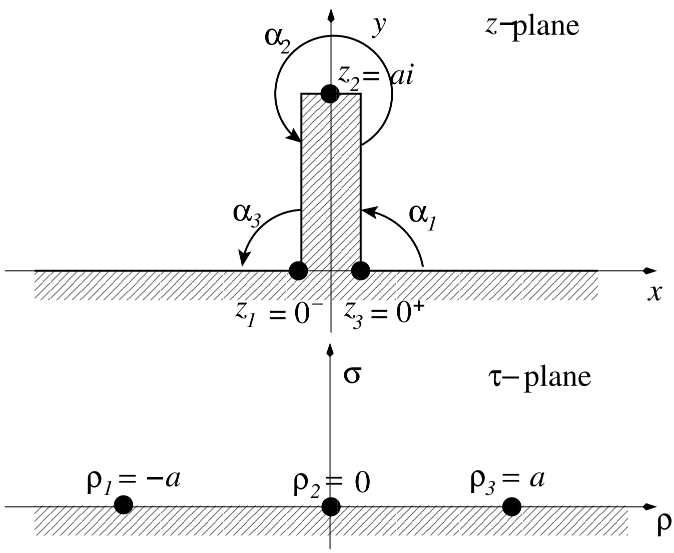


Figure 10.11.4: The conformal mapping between the z-plane and τ-plane achieved by the conformal mapping $\tau = \sqrt{z^2 + a^2}$.

In summary, conformal mapping allowed us to transform the original domain into one (an upper half-plane) where we could construct another analytic function whose imaginary part satisfied Laplace's equation and the boundary conditions. A natural question is what do we do if we cannot find this analytic function in the τ-plane? The next example shows an alternative approach.

• Example 10.11.2

For our second example of conformal mapping, consider $\tau = \sqrt{z^2 + a^2}$. To illustrate this mapping we have constructed two Argand diagrams; one is for the z-plane while the second is for the τ-plane. Figure 10.11.4 shows how a particular boundary in the z-plane maps into the τ-plane. The advantage here is that the infinitely thin filament or peg located at $z = 0$ is completely eliminated in the τ-plane.

One source of concern is the presence of the square root; for any value of z we would have two possible solutions. We make the mapping unique by requiring that $\Im(\tau) \geq 0$.

To better understand this transformation, Figure 10.11.5 illustrates various lines of constant $\Re(\tau/a)$ and $\Im(\tau/a)$ as a function of x/a and y/a. This figure was constructed using the MATLAB code:

```
clear;

% compute τ for various values of z

for jj = 1:40
  y = 0.05 * jj;
for ii = 1:42
  x = 0.05 * (ii-21.5); z = x + i*y; tau(ii,jj) = sqrt(z*z+a*a);
  if (imag(tau(ii,jj)) <= 0) tau(ii,jj) = -tau(ii,jj); end
  X(ii,jj) = x; Y(ii,jj) = y;
  IM(ii,jj) = imag(tau(ii,jj)); REAL(ii,jj) = real(tau(ii,jj));
```

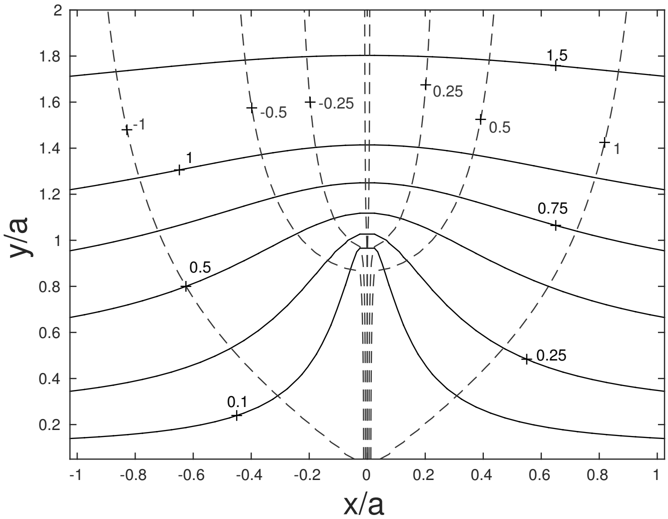


Figure 10.11.5: Lines of constant $\Re(\tau/a)$ (dashed line) and $\Im(\tau/a)$ (solid lines) as a function of x and y for the conformal mapping $\tau = \sqrt{z^2 + a^2}$.

```
end; end
```

% plot the conformal mapping Equation $\tau = \sqrt{z^2 + a^2}$

```
figure
[C,h] = contour(X,Y,IM,[0.1,0.25,0.5,0.75,1,1.5,2],'k');
clabel(C,'FontSize',10,'Color','k','Rotation',0)
xlabel('x','FontSize',20); ylabel('y','FontSize',20);
hold on
v = [-1,-0.5,-0.25,-0.01,0.01,0.25,0.5,1];
[C,h] = contour(X,Y,REAL,v,'--b');
clabel(C,'manual','FontSize',10,'Color','b','Rotation',0)
```

As $y \to \infty$, lines of constant $\Im(\tau/a)$ become parallel to the boundary $y = 0$. Only for smaller values of y and as we approach the peg at $x = 0$ do these lines deviate strongly from the horizontal as they pass over the obstacle. The smaller the value of $\Im(\tau/a)$ the more they conform to the shape of the obstacle.

The behavior of lines of constant $\Re(\tau/a)$ are more difficult to understand. There are two general classes, depending upon whether the absolute value of $\Re(\tau/a)$ is less or greater than 1. When $|\Re(\tau/a)| > 1$ they are clearly orthogonal to constant lines of $\Im(\tau/a)$. Positive values of $\Re(\tau/a)$ exist for $x > 0$ while negative values occur when $x < 0$. $|\Re(\tau/a)| < 1$ for $y \geq a$.

This example has two interesting aspects to it. The first is the presence of the square root. The second involves how we will find the solution to Laplace's equation in the τ-plane.

Let us assume that in the original z-plane the solution equals to zero along the entire boundary except along the "peg." There, the solution equals 1. In the τ-plane the solution equals zero along the entire boundary *except* for the segment $-a < \rho < a$, where $\sigma = 0$, along which the solution equals 1. Instead of finding an analytic function whose real or imaginary part satisfies this boundary condition, we employ the results from Section 11.7.

In the present case, we find that

$$u(\rho,\sigma) = \frac{1}{\pi}\int_{-a}^{a} \frac{\sigma}{\sigma^2+(\xi-\rho)^2}\,d\xi \tag{10.11.4}$$

$$= \frac{1}{\pi}\left[\tan^{-1}\left(\frac{a-\rho}{\sigma}\right)+\tan^{-1}\left(\frac{a+\rho}{\sigma}\right)\right]. \tag{10.11.5}$$

Given Equation 10.11.5 we can compute the solution as follows: For a specific value of x and y, we find the corresponding value of ρ and σ. Equation 10.11.5 gives us the solution to Laplace's equation at that point and the corresponding x and y. The MATLAB code is:

```
clear; a = 1;

for jj = 1:100
  y = 0.02 * jj;
for ii = 1:202
  x = 0.02 * (ii-101.5); z = x + i*y; tau = sqrt(z*z+a*a);
  if (imag(tau) <= 0) tau = -tau; end
  sigma = imag(tau); rho = real(tau);
  X(ii,jj) = x; Y(ii,jj) = y;
% Equation 10.11.5
  arg1 = (a-rho)/sigma; arg2 = (a+rho)/sigma;
  T(ii,jj) = (atan(arg1)+atan(arg2)) / pi;
end; end

% plot the solution to Laplace's equation

figure
[C,h] = contourf(X,Y,T,[0,0.05,0.2,0.4,0.6,0.8],'k');
colormap autumn
clabel(C,'FontSize',10,'Color','k','Rotation',0)
xlabel('x','FontSize',20); ylabel('y','FontSize',20);
```

Figure 10.11.6 illustrates this solution. □

So far we have not presented a strategy for finding our conformal mappings. One method would be to simply experiment with transforms that had been used in similar problems. Fortunately, during the 1860s two German mathematicians, E. B. Christoffel[16] and H. A. Schwarz,[17] developed a very popular method of mapping a polygon into a half plane. Example 10.11.1 illustrated one of their transforms. Indeed, if we imagine that the boundary of the polygon is constructed from a thin wire, the purpose of the Schwarz-Christoffel transformation is to unbend the corners so that the wire becomes straight.

Our derivation begins by considering a mapping $z = f(\tau)$ where

$$\frac{dz}{d\tau} = C(\tau-\rho_1)^{k_1}(\tau-\rho_2)^{k_2}\cdots(\tau-\rho_n)^{k_n}, \tag{10.11.6}$$

[16] Christoffel, E. B., 1868: Sul problema delle temperature stazionarie e la rappresentazione di una data superficie. *Ann. Mat. Pura Appl., Series 2*, **1**, 89–103; Christoffel, E. B., 1870: Sopra un problema proposto da Dirichlet. *Ann. Mat. Pura Appl., Series 2*, **4**, 1–9.

[17] Schwarz, H. A., 1868: Über einige Abbildungsaufgaben. *J. Reine Angew. Math.*, **70**, 105–120.

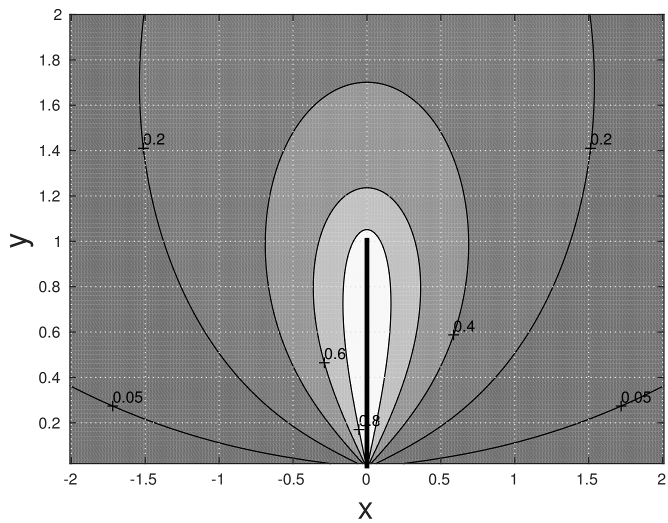


Figure 10.11.6: The solution of Laplace's equation when the solution (potential) along the boundary equals zero except along the peg located at $x = 0$. There the solution (potential) equals one.

and $\rho_1, \rho_2, \ldots, \rho_n$ are any n points arranged in order along the real axis in the τ-plane such that $\rho_1 < \rho_2 < \ldots < \rho_n$. Here the k_i's are real constants and C is a real or complex constant. By taking the logarithm of both sides of Equation 10.11.6 we find that

$$\log\left(\frac{dz}{d\tau}\right) = \log(C) + k_1 \log(\tau - \rho_1) + k_2 \log(\tau - \rho_2) + \cdots + k_n \log(\tau - \rho_n). \qquad \textbf{(10.11.7)}$$

We have assumed that the principal value[18] of each logarithm is taken. The local magnification factor of the mapping from the τ-plane to the z-plane equals $dz/d\tau$ while the angle of $dz/d\tau$ gives the angle through which a small portion of the mapped curve in the τ-plane is rotated by the mapping. This angle is given by

$$\measuredangle\left(\frac{dz}{d\tau}\right) = \measuredangle(C) + k_1 \measuredangle(\tau - \rho_1) + k_2 \measuredangle(\tau - \rho_2) + \cdots + k_n \measuredangle(\tau - \rho_n). \qquad \textbf{(10.11.8)}$$

Equation 10.11.8 follows by first taking the imaginary part of Equation 10.11.7 and then noting that $\measuredangle(C) = \Im[\log(C)]$.

Let the point $(\rho, \sigma) = (-\infty, 0)$ in the τ-plane be mapped into the point z^* in the z-plane. See Figure 10.11.7. If we consider the image of a point ρ as it moves to the right along the negative real axis in the τ-plane, then all of the $\rho - \rho_i$ are real and negative as long as $\rho < \rho_1$. Hence the angles for all of the $\rho - \rho_i$ are constant and equal to π in Equation 10.11.8. Therefore, this equation simplifies to

$$\measuredangle\left(\frac{dz}{d\tau}\right) = \measuredangle(C) + (k_1 + k_2 + \cdots + k_n)\pi. \qquad \textbf{(10.11.9)}$$

Thus the portion of the ρ axis to the left of the point ρ_1 is mapped into a straight line segment, making the angle defined by Equation 10.11.9 with the real axis in the z-plane, and extending from z^* to z_1 the image of $\rho - \rho_1$.

[18] For the complex number $z = re^{\theta i}$, $r \neq 0$, the principle value of the logarithm is $\log(z) = \ln(r) + \theta i$, where θ must lie between 0 and 2π.

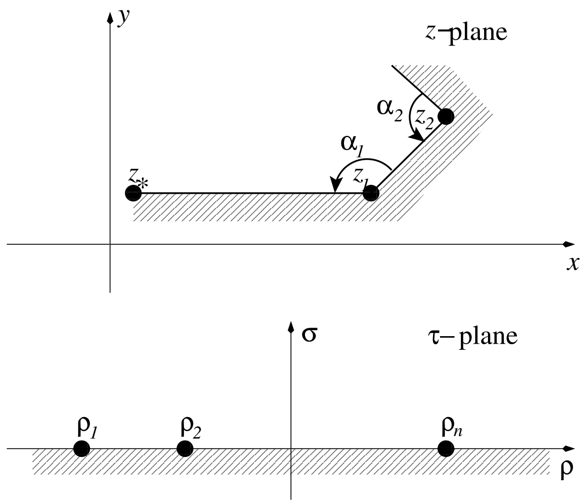


Figure 10.11.7: Diagram used in the derivation of the Schwarz-Christoffel method.

Now as the point ρ crosses the point ρ_1 on the real axis, the real number $\rho - \rho_1$ becomes positive so that its angle abruptly changes from π to 0. Hence $\measuredangle(dz/d\tau)$ abruptly decreases by an amount $k_1\pi$ and then remains constant as τ travels from ρ_1 to ρ_2. It follows that the image of the segment $(\rho_1\rho_2)$ in the z-plane makes an angle of $-k_1\pi$ with the segment (z^*z_1).

Proceeding in this way, we see that each segment (ρ_n, ρ_{n+1}) is mapped into a line segment (z_n, z_{n+1}) in the z-plane, making the angle of $-k_n\pi$ with the segment previously mapped. Thus, if the interior angle of the resultant polynomial contour at the point z_n is to have the magnitude α_n, we must set $\pi - \alpha_n = -k_n\pi$, or $k_n = \alpha_n/\pi - 1$ in Equation 10.11.6. After an integration, we then conclude that the mapping

$$z = C\int^{\tau} (\eta - \rho_1)^{k_1}(\eta - \rho_2)^{k_2} \cdots (\eta - \rho_n)^{k_n}\, d\eta + K, \qquad \textbf{(10.11.10)}$$

where the arbitrary complex constants C and K map the real axis $\sigma = 0$ of the τ-plane into a polynomial boundary in the z-plane in such a way that the vertices $z_1, z_2, \ldots, z_n$ with interior angles $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the images of the points $\rho_1, \rho_2, \ldots, \rho_n$.

For the final segment $\tau - \rho > \rho_n$ the numbers $\tau - \rho_i$ are all real, positive, and equal to zero, so that this segment is rotated through the angle

$$\measuredangle(dz/d\tau) = \measuredangle(C), \qquad \rho > \rho_n. \qquad \textbf{(10.11.11)}$$

For a closed polynomial the sum of the interior angles is

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = (n-2)\pi. \qquad \textbf{(10.11.12)}$$

Therefore,

$$k_1 + k_2 + \cdots + k_n = \frac{(n-2)\pi}{\pi} - n = -2. \qquad \textbf{(10.11.13)}$$

Thus, according to Equations 10.11.8 and 10.11.11, the two infinite segments of the line $\sigma = 0$ are rotated through the angle $\measuredangle(C) - 2\pi$ and $\measuredangle(C)$, as is clearly necessary for a closed figure.

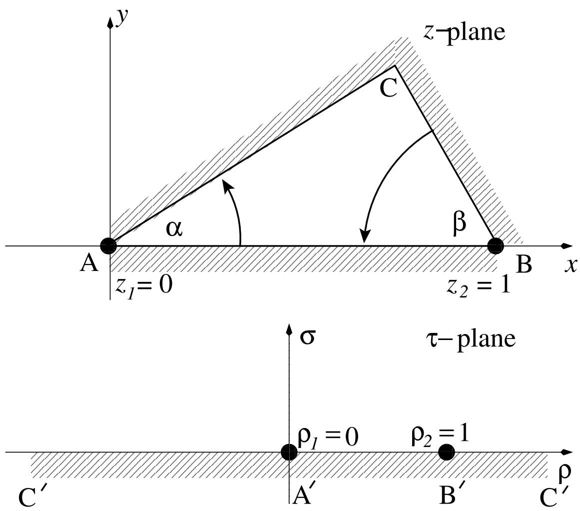


Figure 10.11.8: The complex z- and τ-planes used in Example 10.11.4.

What roles do C and K play? Because C is often complex, this constant introduces any necessary magnification and rotation of the transformation so that *any* prescribed polynomial in the z-plane is made to correspond point by point to the real axis $\sigma = 0$ in the τ-plane. In fact, this correspondence can be set up in infinitely many ways, in that three of the numbers $\rho_1, \rho_2, \ldots, \rho_n$ can be determined arbitrarily. Finally, the mapping can be shown to establish a one-to-one correspondence between points in the interior of the polygon in the z-plane and points in the *upper half* of the τ-plane.

• Example 10.11.3

Let us derive the conformal mapping used in Example 10.11.2. Referring back to Figure 10.11.4, we see that $\alpha_1 = \pi/2$, $k_1 = -1/2$, and $\rho_1 = -a$ at $z_1 = 0^-$; $\alpha_2 = 2\pi$, $k_2 = 1$, and $\rho_2 = 0$ at $z_2 = ai$; and $\alpha_3 = \pi/2$, $k_3 = -1/2$, and $\rho_3 = a$ at $z_3 = 0^+$. Therefore, from Equation 10.11.6,

$$\frac{dz}{d\tau} = C(\tau + a)^{-1/2}\tau(\tau - a)^{-1/2} = C\frac{\tau}{\sqrt{\tau^2 - a^2}}. \qquad \textbf{(10.11.14)}$$

Integrating this differential equation,

$$z = C\sqrt{\tau^2 - a^2} + K. \qquad \textbf{(10.11.15)}$$

Because the point $\rho_1 = -a$ corresponds to $z = 0^-$, $K = 0$. Similarly, at $\rho_2 = 0$, we have that

$$ai = C\sqrt{-a^2}, \qquad \text{or} \qquad C = 1. \qquad \textbf{(10.11.16)}$$

Therefore, the conformal mapping is given by $z = \sqrt{\tau^2 - a^2}$, or $\tau = \sqrt{z^2 + a^2}$. □

• Example 10.11.4

Consider the triangle ABC located in the z-plane as shown on Figure 10.11.8. Here we desire to map the *interior* space of this triangle into the upper half of the τ-plane. At

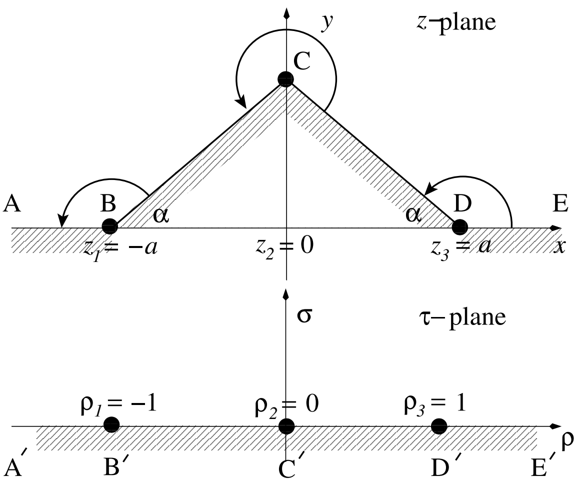


Figure 10.11.9: The complex z- and τ-planes used in Example 10.11.5.

point C, points along the boundary and to the left of C are to be mapped out to $-\infty$ in the τ-plane while points along the boundary and to the right of C are mapped to $+\infty$.

From Equation 10.11.6 we have that

$$\frac{dz}{d\tau} = C'\tau^{\alpha/\pi-1}(\tau-1)^{\beta/\pi-1} = C\tau^{\alpha/\pi-1}(1-\tau)^{\beta/\pi-1}. \qquad \mathbf{(10.11.17)}$$

Integrating this differential equation,

$$z = C\int^{\tau} \eta^{\alpha/\pi-1}(1-\eta)^{\beta/\pi-1}\,d\eta + K. \qquad \mathbf{(10.11.18)}$$

Because we want the points $\tau = 0$ and $z = 0$ to correspond to each other, $K = 0$. On the other hand, if we wish $\tau = 1$ and $z = 1$ to correspond, Equation 10.11.18 yields

$$C\int_0^1 \eta^{\alpha/\pi-1}(1-\eta)^{\beta/\pi-1}\,d\eta = C\,\frac{\Gamma(\alpha/\pi)\Gamma(\beta/\pi)}{\Gamma[(\alpha+\beta)/\pi]} = 1, \qquad \mathbf{(10.11.19)}$$

where $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}\,dt. \qquad \mathbf{(10.11.20)}$$

Consequently,

$$C = \frac{\Gamma[(\alpha+\beta)/\pi]}{\Gamma(\alpha/\pi)\Gamma(\beta/\pi)}, \qquad \mathbf{(10.11.21)}$$

and

$$z = \frac{\Gamma[(\alpha+\beta)/\pi]}{\Gamma(\alpha/\pi)\Gamma(\beta/\pi)}\int_0^{\tau} \eta^{\alpha/\pi-1}(1-\eta)^{\beta/\pi-1}\,d\eta. \qquad \mathbf{(10.11.22)}$$

A noteworthy aspect of this example is that the conformal mapping is given by an integral and not some analytic expression. □

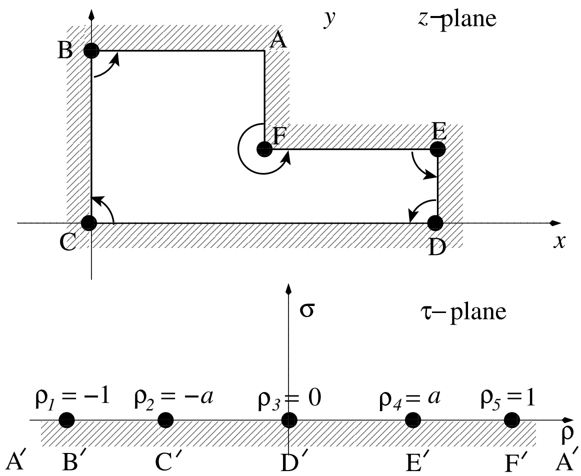


Figure 10.11.10: The complex z- and τ-planes used in Example 10.11.6 with $a < 1$.

• Example 10.11.5

Consider the domain lying in the upper half of the z-plane except for a triangular section BCD shown in Figure 10.11.9. We wish to construct the Schwarz-Christoffel transformation that maps this domain into the upper half of the τ-plane. From Equation 10.11.6 we have that

$$\frac{dz}{d\tau} = C'(\tau+1)^{(\pi-\alpha)/\pi-1}\tau^{(\pi+2\alpha)/\pi-1}(\tau-1)^{(\pi-\alpha)/\pi-1} \quad (\mathbf{10.11.23})$$

$$= C'\frac{\tau^{2\alpha/\pi}}{(\tau^2-1)^{\alpha/\pi}} = C\frac{\tau^{2\alpha/\pi}}{(1-\tau^2)^{\alpha/\pi}}. \quad (\mathbf{10.11.24})$$

Integrating this differential equation,

$$z = C\int_0^\tau \frac{\eta^{2\alpha/\pi}}{(1-\eta^2)^{\alpha/\pi}}\,d\eta + K. \quad (\mathbf{10.11.25})$$

If we want the point $\tau = 0$ to correspond to the point $z = ki$, then $K = ki$. On the other hand, if the point $\tau = 1$ corresponds to $z = a$, then

$$a = C\int_0^1 \frac{\eta^{2\alpha/\pi}}{(1-\eta^2)^{\alpha/\pi}}\,d\eta + ki. \quad (\mathbf{10.11.26})$$

Solving for C,

$$C = \frac{\sqrt{\pi}(a-ki)}{\Gamma\left(\alpha/\pi+\frac{1}{2}\right)\Gamma(1-\alpha/\pi)}. \quad (\mathbf{10.11.27})$$

Therefore, the final answer is

$$z = \frac{\sqrt{\pi}(a-ki)}{\Gamma\left(\alpha/\pi+\frac{1}{2}\right)\Gamma(1-\alpha/\pi)}\int_0^\tau \frac{\eta^{2\alpha/\pi}}{(1-\eta^2)^{\alpha/\pi}}\,d\eta + ki. \quad (\mathbf{10.11.28})$$

□

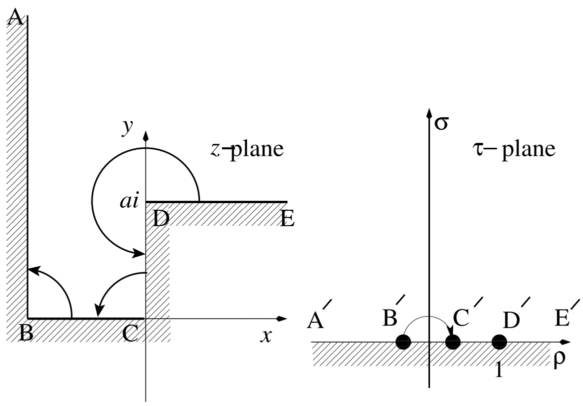


Figure 10.11.11: The complex z- and τ-planes used in Example 10.11.7.

• Example 10.11.6

Consider the domain within the L-shaped boundary shown in Figure 10.11.10. We wish to construct the Schwarz-Christoffel transform that maps the interior into the upper half of the τ-plane. Note that we broke the boundary in such a manner that points slightly to the left of point A are mapped to $-\infty$ while points slightly below the point A are mapped to $+\infty$.

Because $a < 1$, Equation 10.11.6 gives

$$\frac{dz}{d\tau} = C(\tau+1)^{-1/2}(\tau+a)^{-1/2}\tau^{-1/2}(\tau-a)^{-1/2}(\tau-1)^{1/2}. \tag{10.11.29}$$

Integrating this differential equation,

$$z = C\int_0^\tau \frac{(\eta-1)\,d\eta}{\eta\sqrt{(\eta^2-1)(\eta^2-a^2)}} + K = \frac{C}{a}\int_0^\tau \frac{(\eta-1)\,d\eta}{\eta\sqrt{(1-\eta^2)(1-p^2\eta^2)}} + K, \tag{10.11.30}$$

where $p^2 = 1/a^2$. To compute C and K, we would need further information.

□

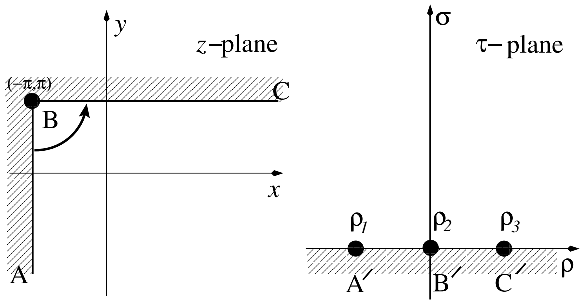
• Example 10.11.7

Let us derive the conformal mapping, Equation 10.11.2, used in Example 10.11.1. The $z-$ and $\tau-$planes are shown in Figure 10.11.11. From this figure we see that $\alpha_1 = 3\pi/2$, $\alpha_2 = \pi/2$, $\alpha_3 = \pi/2$, $\rho_1 = 1$, $\rho_2 = 0^-$, and $\rho_3 = 0^+$. This yields

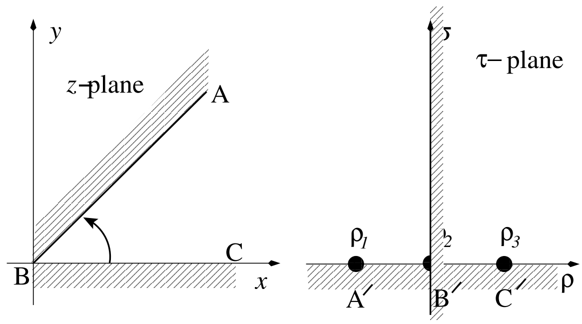
$$\frac{dz}{d\tau} = K(\tau-1)^{(3\pi)/(2\pi)-1}(\tau-0^-)^{(\pi)/(2\pi)-1}(\tau-0^+)^{(\pi)/(2\pi)-1} = K\frac{\sqrt{\tau-1}}{\tau}. \tag{10.11.31}$$

Integrating Equation 10.11.31, we find that

$$z = 2K\left[\sqrt{\tau-1} - \arctan\left(\sqrt{\tau-1}\,\right)\right] + C = 2K\left[\sqrt{\tau-1} + \frac{i}{2}\log\left(\frac{1+i\sqrt{\tau-1}}{1-i\sqrt{\tau-1}}\right)\right] + C. \tag{10.11.32}$$



Problem 3



Problem 4

Because at $\tau = 1$, $z = a/2$, we have $C = a/2$.

The computation of K is more complicated. Referring to Figure 10.11.11, we note that

$$\int_B^C dz = \int_{B'}^{C'} K \frac{\sqrt{\tau - 1}}{\tau}\, d\tau. \qquad \textbf{(10.11.33)}$$

Setting $\tau = r\, e^{\theta i}$ with $r \to 0$, Equation 10.11.34 becomes

$$\frac{a}{2} = K \lim_{r\to 0} \int_\pi^0 \frac{\sqrt{r\, e^{\theta i} - 1}}{r\, e^{\theta i}}\, ir\, e^{\theta i}\, d\theta = K\pi. \qquad \textbf{(10.11.34)}$$

Thus $K = a/(2\pi)$ and we recover Equation 10.11.2.

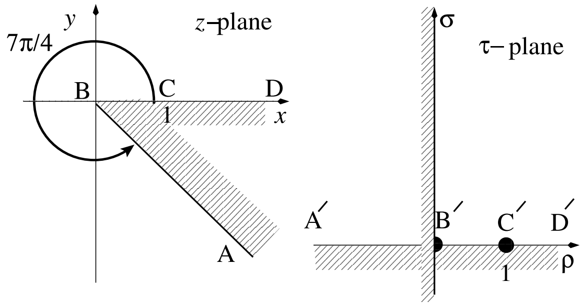
Problems

1. Verify that the function $\tau = e^z$ maps the strip $0 < \Im(z) < \pi$ into the half-plane $\Im(\tau) > 0$.

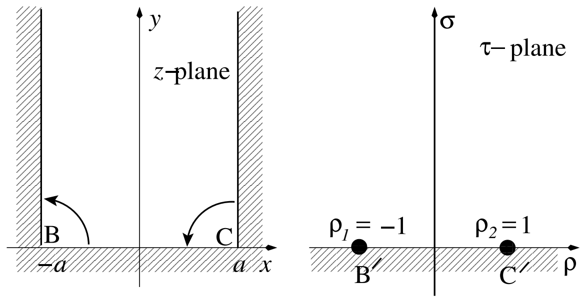
2. Verify that the function $\tau^2 = 1 - e^z$ maps the strip $-\pi < \Im(z) < \pi$, except for the negative real axis, into the upper half of the τ-plane.

3. Use the Schwarz-Christoffel method to find the conformal mapping that maps the quarter plane $x > -\pi$, $y < \pi$ into the upper half of the τ-plane. We require that the point $(-\pi, \pi)$ in the z-plane maps to the point $(0, 0)$ in the τ-plane.

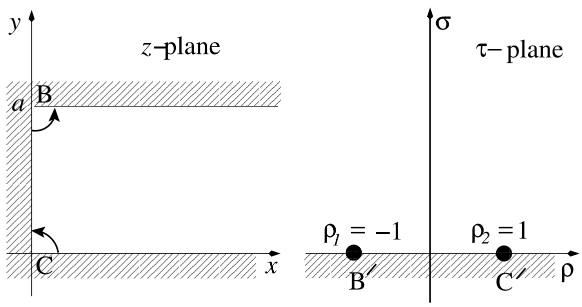
4. Use the Schwarz-Christoffel method to find the conformal mapping that maps the sector lying between the x-axis and the line $\theta = \pi/3$ into the upper half of the τ-plane. We require that the point $(0, 0)$ in the z-plane maps to the point $(0, 0)$ in the τ-plane.



Problem 5



Problem 6



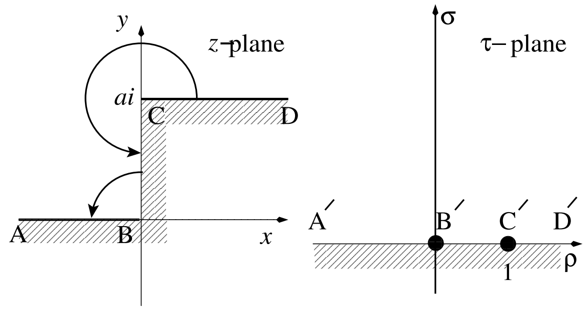
Problem 7

5. Use the Schwarz-Christoffel method to find the conformal mapping that maps the portion of the z-plane defined by $0 < r < \infty$, $0 < \theta < 7\pi/4$ into the upper half of the τ-plane. We require that the points $(0, 0)$ and $(1, 0)$ in the z-plane map to the points $(0, 0)$ and $(1, 0)$ in the τ-plane, respectively.

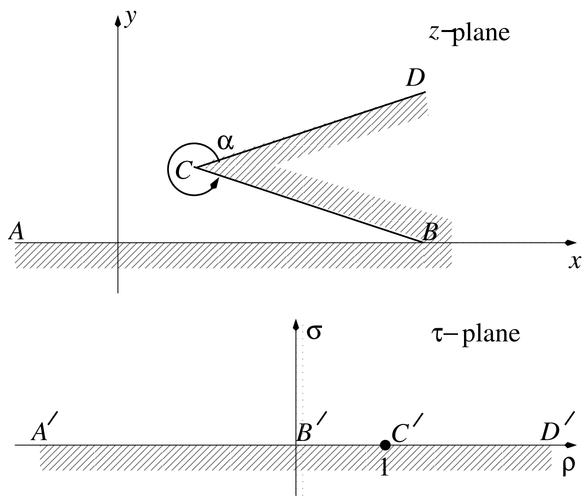
6. Use the Schwarz-Christoffel method to find the conformal mapping that maps the domain $|x| < a$, $0 < y$ into the upper half of the τ-plane. Let the point $(-a, 0)$ become the point $(-1, 0)$ while the point $(a, 0)$ becomes the point $(1, 0)$.

7. Use the Schwarz-Christoffel method to find the conformal mapping that maps the region $x > 0$, $0 < y < a$ into the upper half of the τ-plane. We require that the point $(0, a)$ maps to $(-1, 0)$ in the τ-plane while the point $(0, 0)$ maps to $(1, 0)$ in the τ-plane.

8. Use the Schwarz-Christoffel method to find the conformal mapping that maps the region shown in the figure into the upper half of the τ-plane. We require that the points $(0, 0)$ and $(0, a)$ in the z-plane map to the points $(0, 0)$ and $(1, 0)$ in the τ-plane, respectively.



Problem 8



Problem 9a

9. Referring to the figure entitled Problem 9b, construct a transform between a z-plane which has a barrier that runs parallel to the x-axis from $z = L + \pi L\,i$ to $\infty + \pi L\,i$ and a τ-plane that has no barrier.

Step 1: Referring to the figure entitled Problem 9a, use the Schwarz-Christoffel method to show that the transform is given by

$$\frac{dz}{d\tau} = C\tau^{k_1}(\tau - 1)^{k_2},$$

where $k_1 = -\alpha/(2\pi)$ and $k_2 = \alpha/\pi - 1$.

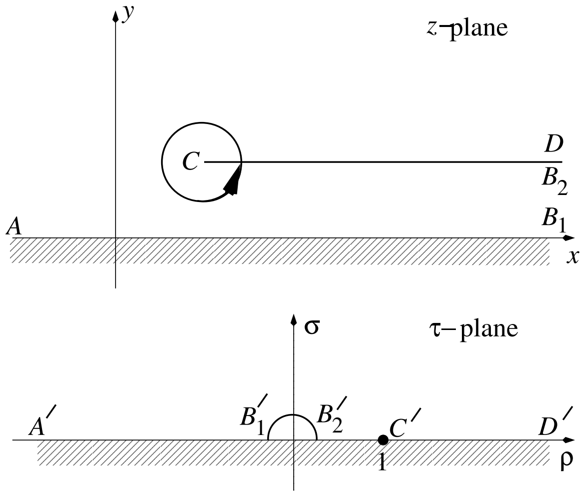
Step 2: Consider the situation as the points B and D in the z-plane move out to infinity (so that $\alpha \to 2\pi$). Show that the transform approaches

$$\frac{dz}{d\tau} = C\frac{\tau - 1}{\tau},$$

or

$$z = C[\tau - \log(\tau)] + K.$$

Here we have taken the principal branch of the logarithm so that $\log(z) = \ln(|z|) + i\theta$ where $0 \leq \theta \leq \pi$. (We do not require that $0 \leq \theta < 2\pi$ because we are always in the upper half-plane.) See the figure entitled Problem 9b.



Problem 9b

Step 3: Following Example 10.11.7, consider the area around $\tau = 0$. Show that

$$dz \approx -C\frac{d\tau}{\tau} = -iC\,d\theta,$$

where $\tau = r\,e^{\theta i}$. Integrating from point B_1' to point B_2', show that $C = L$.

Step 4: To compute K, note that if the point C, located at $z = L + \pi Li$, corresponds to the point C', located at $\tau = 1$, then $K = \pi Li$.

10. Use conformal mapping to solve Laplace's equation for the infinite strip $-\infty < x < \infty$, $0 \le y \le \pi$. The solution equals zero everywhere along the boundary *except* for $x > 0$, $y = 0$, where $u(x, 0) = 1$.

Step 1: Consider the mapping $\tau = e^z$. Show that $\rho = e^x\cos(y)$ and $\sigma = e^x\sin(y)$. In particular, $(\infty, \pi) \to (-\infty, 0)$, $(0, \pi) \to (-1, 0)$, $(-\infty, y) \to (0, 0)$, $(0, 0) \to (1, 0)$, and $(\infty, 0) \to (\infty, 0)$.

Step 2: Using the Fourier method from Section 11.7, show that

$$u(\rho, \sigma) = \frac{1}{\pi}\left[\frac{\pi}{2} - \tan^{-1}\left(\frac{1-x}{y}\right)\right] = 1 - \frac{1}{\pi}\tan^{-1}\left(\frac{y}{x-1}\right).$$

Step 3: Show that

$$u(x, y) = 1 - \frac{1}{\pi}\tan^{-1}\left[\frac{e^x\sin(y)}{e^x\cos(y) - 1}\right].$$

11. Use conformal mapping to solve Laplace's equation for a pie-shaped sector in the first quadrant. See Figure 10.11.12. The solution equals zero along the entire boundary except for $0 < x < 1$ where it equals one.

Step 1: Show that the mapping $z = \tau^{\alpha/\pi}$ or $\tau = z^{\pi/\alpha}$ maps the pie-shaped sector into the half-plane $\Im(\tau) > 0$. See Figure 10.11.12.

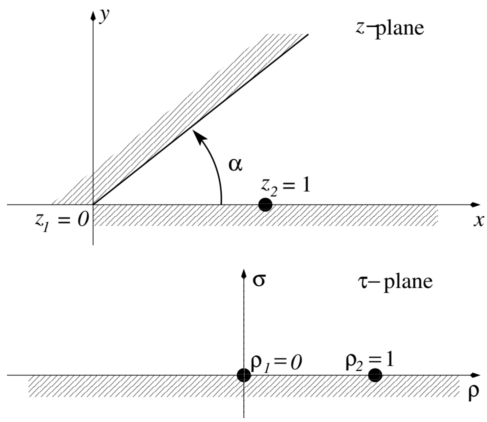


Figure 10.11.12: The conformal mapping between the z-plane and τ-plane achieved by the conformal mapping $\tau = z^{\pi/\alpha}$.

Step 2: Using the Fourier method from Section 11.7, show that

$$u(\rho, \sigma) = \frac{1}{\pi} \cot^{-1}\left(\frac{\rho^2 + \sigma^2 - \rho}{\sigma} \right).$$

Step 3: Show that

$$u(r, \theta) = \frac{1}{\pi} \cot^{-1}\left[\frac{r^{\pi/\alpha} - \cos(\pi\theta/\alpha)}{\sin(\pi\theta/\alpha)} \right],$$

where $x = r\cos(\theta)$ and $y = r\sin(\theta)$.

12. Use conformal mapping to solve Laplace's equation for the semi-infinite strip $0 \le x \le a$, $0 \le y < \infty$, where $u(x, 0) = 1$, $0 \le x \le a$, and $u(0, y) = u(a, y) = 0$, $0 \le y < \infty$.

Step 1: Consider the mapping $\tau = -\cos(\pi z/a)$. Show that

$$\rho = -\cos(\pi x/a)\cosh(\pi y/a), \quad \text{and} \quad \sigma = \sin(\pi x/a)\sinh(\pi y/a).$$

In particular, $(0, \infty) \to (-\infty, 0)$, $(0, 0) \to (-1, 0)$, $(a/2, 0) \to (0, 0)$, $(a, 0) \to (1, 0)$, and $(a, \infty) \to (\infty, 0)$.

Step 2: Using the Fourier method from Section 11.7, show that

$$u(\rho, \sigma) = \frac{1}{\pi} \cot^{-1}\left(\frac{\rho^2 + \sigma^2 - 1}{2\sigma} \right).$$

Step 3: Show that

$$\begin{aligned} u(x, y) &= \frac{1}{\pi} \cot^{-1}\left\{ \left[\sinh^2\left(\frac{\pi y}{a}\right) - \sin^2\left(\frac{\pi x}{a}\right) \right] \Big/ 2\sin\left(\frac{\pi x}{a}\right) \sinh\left(\frac{\pi y}{a}\right) \right\} \\ &= \frac{2}{\pi} \tan^{-1}\left[\sin\left(\frac{\pi x}{a}\right) \Big/ \sinh\left(\frac{\pi y}{a}\right) \right]. \end{aligned}$$

Step 4: In the case that boundary conditions read $u(0, y) = u(a, y) = 1$ for $0 \le y < \infty$ and $u(x, 0) = 0$ for $0 \le x \le a$, how could you use the solution in Step 3 to solve this new problem?

Further Readings

Ablowitz, M. J., and A. S. Fokas, 2003: *Complex Variables: Introduction and Applications*. Cambridge University Press, 660 pp. Covers a wide variety of topics, including complex numbers, analytic functions, singularities, conformal mapping and the Riemann-Hilbert problem.

Carrier, G. F., M. Krook, and C. E. Pearson, 1966: *Functions of a Complex Variable: Theory and Technique*. McGraw-Hill Book Co., 438 pp. Graduate-level textbook.

Churchill, R. V., 1960: *Complex Variables and Applications*. McGraw-Hill Book Co., 297 pp. Classic textbook.

Flanigan, F. J., 1983: *Complex Variables*. Dover, 364 pp. A crystal clear exposition and emphasis on an intuitive understanding of complex analysis.

Chapter 11
The Fourier Transform

In the previous chapter we showed how we could expand a periodic function in terms of an infinite sum of sines and cosines. However, most functions encountered in engineering are aperiodic. As we shall see, the extension of Fourier series to these functions leads to the Fourier transform.

11.1 FOURIER TRANSFORMS

The Fourier transform is the natural extension of Fourier series to a function $f(t)$ of infinite period. To show this, consider a periodic function $f(t)$ of period $2T$ that satisfies the so-called Dirichlet's conditions.[1] If the integral $\int_a^b |f(t)|\,dt$ exists, this function has the complex Fourier series

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi t/T}, \tag{11.1.1}$$

where

$$c_n = \frac{1}{2T} \int_{-T}^{T} f(t) e^{-in\pi t/T} dt. \tag{11.1.2}$$

Equation 11.1.1 applies only if $f(t)$ is continuous at t; if $f(t)$ suffers from a jump discontinuity at t, then the left side of Equation 11.1.1 equals $\frac{1}{2}[f(t^+) + f(t^-)]$, where $f(t^+) = \lim_{x\to t^+} f(x)$ and $f(t^-) = \lim_{x\to t^-} f(x)$. Substituting Equation 11.1.2 into Equation 11.1.1,

$$f(t) = \frac{1}{2T} \sum_{n=-\infty}^{\infty} e^{in\pi t/T} \int_{-T}^{T} f(x) e^{-in\pi x/T} dx. \tag{11.1.3}$$

[1] A function $f(t)$ satisfies Dirichlet's conditions in the interval (a, b) if (1) it is bounded in (a, b), and (2) it has at most a finite number of discontinuities and a finite number of maxima and minima in that interval.

Let us now introduce the notation $\omega_n = n\pi/T$ so that $\Delta\omega_n = \omega_{n+1} - \omega_n = \pi/T$. Then,

$$f(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(\omega_n) e^{i\omega_n t} \Delta\omega_n, \tag{11.1.4}$$

where

$$F(\omega_n) = \int_{-T}^{T} f(x) e^{-i\omega_n x} \, dx. \tag{11.1.5}$$

As $T \to \infty$, ω_n approaches a continuous variable ω, and $\Delta\omega_n$ may be interpreted as the infinitesimal $d\omega$. Therefore, ignoring any possible difficulties,[2]

$$\boxed{f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \, d\omega,} \tag{11.1.6}$$

and

$$\boxed{F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \, dt.} \tag{11.1.7}$$

Equation 11.1.7 is the *Fourier transform* of $f(t)$ while Equation 11.1.6 is the *inverse Fourier transform* that converts a Fourier transform back to $f(t)$. Alternatively, we may combine Equation 11.1.6 and Equation 11.1.7 to yield the equivalent real form

$$f(t) = \frac{1}{\pi} \int_0^{\infty} \left\{ \int_{-\infty}^{\infty} f(x) \cos[\omega(t-x)] \, dx \right\} d\omega. \tag{11.1.8}$$

Hamming[3] suggested the following analog in understanding the Fourier transform. Let us imagine that $f(t)$ is a light beam. Then the Fourier transform, like a glass prism, breaks up the function into its component frequencies ω, each of intensity $F(\omega)$. In optics, the various frequencies are called colors; by analogy the Fourier transform gives us the color spectrum of a function. On the other hand, the inverse Fourier transform blends a function's spectrum to give back the original function.

Most signals encountered in practice have Fourier transforms because they are absolutely integrable, since they are bounded and of finite duration. However, there are some notable exceptions. Examples include the trigonometric functions sine and cosine.

---

[2] For a rigorous derivation, see Titchmarsh, E. C., 1948: *Introduction to the Theory of Fourier Integrals*. Oxford University Press, Chapter 1.

[3] Hamming, R. W., 1977: *Digital Filters*. Prentice-Hall, p. 136.

• Example 11.1.1

Let us find the Fourier transform for

$$f(t) = \begin{cases} 1, & |t| < a, \\ 0, & |t| > a. \end{cases} \tag{11.1.9}$$

From the definition of the Fourier transform,

$$F(\omega) = \int_{-\infty}^{-a} 0\, e^{-i\omega t}\, dt + \int_{-a}^{a} 1\, e^{-i\omega t}\, dt + \int_{a}^{\infty} 0\, e^{-i\omega t}\, dt \tag{11.1.10}$$

$$= \frac{e^{\omega a i} - e^{-\omega a i}}{\omega i} = \frac{2\sin(\omega a)}{\omega} = 2a \operatorname{sinc}(\omega a), \tag{11.1.11}$$

where $\operatorname{sinc}(x) = \sin(x)/x$ is the *sinc function.*

Although this particular example does not show it, the Fourier transform is, in general, a complex function. The most common method of displaying it is to plot its amplitude and phase on two separate graphs for all values of ω. Another problem here is ratio of 0/0 when $\omega = 0$. Applying L'Hôpital's rule, we find that $F(0) = 2$. Thus, we can plot the amplitude and phase of $F(\omega)$ using the MATLAB script:

```
clear; % clear all previous computations
omegan = [-20:0.01:-0.01]; % set up negative frequencies
omegap = [0.01:0.01:20]; % set up positive frequencies
% compute Fourier transform for negative frequencies
f_omegan = 2.*sin(omegan)./omegan;
% compute Fourier transform for positive frequencies
f_omegap = 2.*sin(omegap)./omegap;
% concatenate all of the frequencies
omega = [omegan,0,omegap];
% bring together the Fourier transforms found
%     at positive and negative frequencies
f_omega = [f_omegan,2,f_omegap];
amplitude = abs(f_omega); % compute the amplitude
phase = atan2(0,f_omega); % compute the phase
clf; % clear all previous figures
% plot frequency spectrum
subplot(2,1,1), plot(omega,amplitude)
% label amplitude plot
ylabel('|F(\omega)|/a','FontSize',15)
subplot(2,1,2), plot(omega,phase) % plot phase of transform
ylabel('phase','FontSize',15) % label amplitude plot
xlabel('\omega','FontSize',15) % label x-axis.
```

Figure 11.1.1 shows the output from the MATLAB script. Of these two quantities, the amplitude is by far the more popular one and is given the special name of *frequency spectrum.* □

From the definition of the inverse Fourier transform,

$$f(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin(\omega a)}{\omega} e^{i\omega t}\, d\omega = \begin{cases} 1, & |t| < a, \\ 0, & |t| > a. \end{cases} \tag{11.1.12}$$

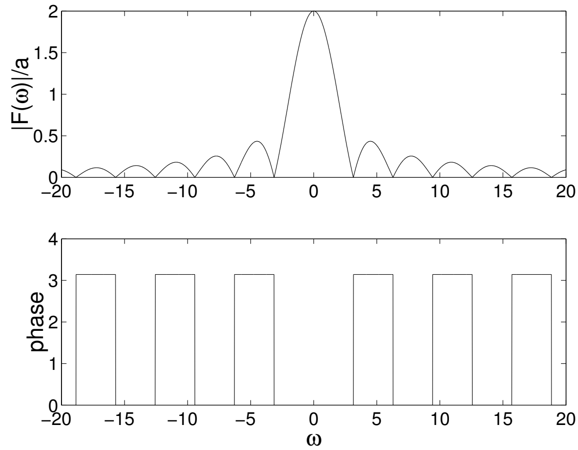


Figure 11.1.1: Graph of the Fourier transform for Equation 11.1.9.

An important question is what value does $f(t)$ converge to in the limit as $t \to a$ and $t \to -a$? Because Fourier transforms are an extension of Fourier series, the behavior at a jump is the same as that for a Fourier series. For that reason, $f(a) = \frac{1}{2}[f(a^+)+f(a^-)] = \frac{1}{2}$ and $f(-a) = \frac{1}{2}[f(-a^+)+f(-a^-)] = \frac{1}{2}$.

• Example 11.1.2: Dirac delta function

Of the many functions that have a Fourier transform, a particularly important one is the *(Dirac) delta function*.[4] For example, in Section 11.6 we will use it to solve differential equations. We *define* it as the inverse of the Fourier transform $F(\omega) = 1$. Therefore,

$$\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\,d\omega. \tag{11.1.13}$$

□

To give some insight into the nature of the delta function, consider another band-limited transform

$$F_\Omega(\omega) = \begin{cases} 1, & |\omega| < \Omega, \\ 0, & |\omega| > \Omega, \end{cases} \tag{11.1.14}$$

where Ω is real and positive. Then,

$$f_\Omega(t) = \frac{1}{2\pi}\int_{-\Omega}^{\Omega} e^{i\omega t}\,d\omega = \frac{\Omega}{\pi}\frac{\sin(\Omega t)}{\Omega t}. \tag{11.1.15}$$

Figure 11.1.2 illustrates $f_\Omega(t)$ for a large value of Ω. We observe that as $\Omega \to \infty$, $f_\Omega(t)$ becomes very large near $t = 0$ as well as very narrow. On the other hand, $f_\Omega(t)$ rapidly approaches zero as $|t|$ increases. Therefore, the delta function is given by the limit

$$\delta(t) = \lim_{\Omega\to\infty} \frac{\sin(\Omega t)}{\pi t}, = \begin{cases} \infty, & t = 0, \\ 0, & t \neq 0. \end{cases} \tag{11.1.16}$$

[4] Dirac, P. A. M., 1947: *The Principles of Quantum Mechanics*. Oxford University Press, Section 15.

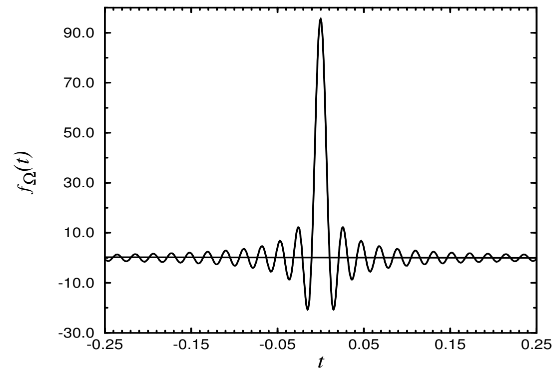


Figure 11.1.2: Graph of the function given in Equation 11.1.15 for $\Omega = 300$.

Because the Fourier transform of the delta function equals one,

$$\int_{-\infty}^{\infty} \delta(t) e^{-i\omega t} dt = 1. \tag{11.1.17}$$

Since Equation 11.1.17 must hold for any ω, we take $\omega = 0$ and find that

$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1. \tag{11.1.18}$$

Thus, the area under the delta function equals unity. Taking Equation 11.1.16 into account, we can also write Equation 11.1.18 as

$$\int_{-a}^{b} \delta(t)\, dt = 1, \qquad a, b > 0. \tag{11.1.19}$$

Finally, from the law of the mean of integrals, we have the *sifting property* that

$$\int_{a}^{b} f(t)\delta(t - t_0)\, dt = f(t_0), \tag{11.1.20}$$

if $a < t_0 < b$. This property is given its name because $\delta(t - t_0)$ acts as a sieve, selecting from all possible values of $f(t)$ its value at $t = t_0$.

We can also use several other functions with equal validity to represent the delta function. These include the limiting case of the following rectangular or triangular distributions:

$$\delta(t) = \lim_{\epsilon \to 0} \begin{cases} \frac{1}{\epsilon}, & |t| < \frac{\epsilon}{2}, \\ 0, & |t| > \frac{\epsilon}{2}, \end{cases} \quad \text{or} \quad \delta(t) = \lim_{\epsilon \to 0} \begin{cases} \frac{1}{\epsilon}\left(1 - \frac{|t|}{\epsilon}\right), & |t| < \epsilon, \\ 0, & |t| > \epsilon, \end{cases} \tag{11.1.21}$$

and the Gaussian function:

$$\delta(t) = \lim_{\epsilon \to 0} \frac{\exp(-\pi t^2/\epsilon)}{\sqrt{\epsilon}}. \tag{11.1.22}$$

Note that the delta function is an even function.

The Fourier Transforms of Some Commonly Encountered Functions

	$\mathbf{f(t)},\ \|\mathbf{t}\| < \infty$	$\mathbf{F}(\omega)$
1.	$e^{-at}H(t), \quad a > 0$	$\dfrac{1}{a+\omega i}$
2.	$e^{at}H(-t), \quad a > 0$	$\dfrac{1}{a-\omega i}$
3.	$te^{-at}H(t), \quad a > 0$	$\dfrac{1}{(a+\omega i)^2}$
4.	$te^{at}H(-t), \quad a > 0$	$\dfrac{-1}{(a-\omega i)^2}$
5.	$t^n e^{-at}H(t),\ \Re(a) > 0,\ n = 1, 2, \ldots$	$\dfrac{n!}{(a+\omega i)^{n+1}}$
6.	$e^{-a\|t\|}, \quad a > 0$	$\dfrac{2a}{\omega^2 + a^2}$
7.	$te^{-a\|t\|}, \quad a > 0$	$\dfrac{-4a\omega i}{(\omega^2 + a^2)^2}$
8.	$\dfrac{1}{1+a^2t^2}$	$\dfrac{\pi}{\|a\|} e^{-\|\omega/a\|}$
9.	$\dfrac{\cos(at)}{1+t^2}$	$\frac{\pi}{2}\left(e^{-\|\omega-a\|} + e^{-\|\omega+a\|}\right)$
10.	$\dfrac{\sin(at)}{1+t^2}$	$\frac{\pi}{2i}\left(e^{-\|\omega-a\|} - e^{-\|\omega+a\|}\right)$
11.	$\begin{cases} 1, & \|t\| < a \\ 0, & \|t\| > a \end{cases}$	$\dfrac{2\sin(\omega a)}{\omega}$
12.	$\dfrac{\sin(at)}{at}$	$\begin{cases} \pi/a, & \|\omega\| < a \\ 0, & \|\omega\| > a \end{cases}$
13.	$e^{-at^2}, \quad a > 0$	$\sqrt{\dfrac{\pi}{a}} \exp\left(-\dfrac{\omega^2}{4a}\right)$

Note: The Heaviside step function $H(t)$ is defined by Equation 11.1.31.

• **Example 11.1.3: Multiple Fourier transforms**

The concept of Fourier transforms can be extended to multivariable functions. Consider a two-dimensional function $f(x, y)$. Then, holding y constant,

$$G(\xi, y) = \int_{-\infty}^{\infty} f(x, y)\, e^{-i\xi x}\, dx. \tag{11.1.23}$$

Then, holding ξ constant,

$$F(\xi, \eta) = \int_{-\infty}^{\infty} G(\xi, y)\, e^{-i\eta y}\, dy. \tag{11.1.24}$$

Therefore, the double Fourier transform of $f(x, y)$ is

$$F(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, e^{-i(\xi x + \eta y)}\, dx\, dy, \tag{11.1.25}$$

assuming that the integral exists.

In a similar manner, we can compute $f(x,y)$ given $F(\xi,\eta)$ by reversing the process. Starting with

$$G(\xi,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\xi,\eta)\, e^{i\eta y}\, d\eta, \qquad (\mathbf{11.1.26})$$

followed by

$$f(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(\xi,y)\, e^{i\xi x}\, d\xi, \qquad (\mathbf{11.1.27})$$

we find that

$$f(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\xi,\eta)\, e^{i(\xi x+\eta y)}\, d\xi\, d\eta. \qquad (\mathbf{11.1.28})$$

□

• **Example 11.1.4: Computation of Fourier transforms using MATLAB**

The Heaviside (unit) step function is a piecewise continuous function defined by

$$H(t-a) = \begin{cases} 1, & t > a, \\ 0, & t < a, \end{cases} \qquad (\mathbf{11.1.29})$$

where $a \ge 0$. We will have much to say about this very useful function in the chapter on Laplace transforms. Presently we will use it to express functions whose definition changes over different ranges of t. For example, the "top hat" function Equation 11.1.9 can be rewritten $f(t) = H(t+a) - H(t-a)$. We can see that this is correct by considering various ranges of t. For example, if $t < -a$, both step functions equal zero and $f(t) = 0$. On the other hand, if $t > a$, both step functions equal one and again $f(t) = 0$. Finally, for $-a < t < a$, the first step function equals one while the second one equals zero. In this case, $f(t) = 1$. Therefore, $f(t) = H(t+a) - H(t-a)$ is equivalent to Equation 11.1.9.

This ability to rewrite functions in terms of the step function is crucial if you want to use MATLAB to compute the Fourier transform via the MATLAB routine `fourier`. For example, how would we compute the Fourier transform of the signum function? The MATLAB commands

```
>> syms omega t; syms a positive
>> fourier('Heaviside(t+a)-Heaviside(t-a)',t,omega)
>> simplify(ans)
```

yields

```
ans =
2*sin(a*omega)/omega
```

the correct answer.

Problems

1. (a) Show that the Fourier transform of

$$f(t) = e^{-a|t|}, \qquad a > 0, \qquad \text{is} \qquad F(\omega) = \frac{2a}{\omega^2 + a^2}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Use MATLAB's `fourier` to find $F(\omega)$.

2. (a) Show that the Fourier transform of

$$f(t) = te^{-a|t|}, \quad a > 0, \qquad \text{is} \qquad F(\omega) = -\frac{4a\omega i}{(\omega^2 + a^2)^2}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Use MATLAB's `fourier` to find $F(\omega)$.

3. (a) Show that the Fourier transform of

$$f(t) = e^{-at^2}, \quad a > 0, \qquad \text{is} \qquad F(\omega) = \sqrt{\frac{\pi}{a}} \exp\left(-\frac{\omega^2}{4a}\right).$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Use MATLAB's `fourier` to find $F(\omega)$.

4. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} e^{2t}, & t < 0, \\ e^{-t}, & t > 0, \end{cases} \qquad \text{is} \qquad F(\omega) = \frac{3}{(2 - i\omega)(1 + i\omega)}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

5. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} e^{-(1+i)t}, & t > 0, \\ -e^{(1-i)t}, & t < 0, \end{cases} \qquad \text{is} \qquad F(\omega) = \frac{-2i(\omega + 1)}{(\omega + 1)^2 + 1}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

6. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} \cos(at), & |t| < 1, \\ 0, & |t| > 1, \end{cases} \qquad \text{is} \qquad F(\omega) = \frac{\sin(\omega - a)}{\omega - a} + \frac{\sin(\omega + a)}{\omega + a}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

7. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} \sin(t), & 0 \le t < 1, \\ 0, & \text{otherwise}, \end{cases}$$

is

$$F(\omega) = -\frac{1}{2}\left[\frac{1-\cos(\omega-1)}{\omega-1} + \frac{\cos(\omega+1)-1}{\omega+1}\right] - \frac{i}{2}\left[\frac{\sin(\omega-1)}{\omega-1} - \frac{\sin(\omega+1)}{\omega+1}\right].$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

8. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} t/a, & |t| < a, \\ 0, & |t| > a, \end{cases} \qquad \text{is} \qquad F(\omega) = \frac{2i\cos(\omega a)}{\omega} - \frac{2i\sin(\omega a)}{\omega^2 a}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

9. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} (t/a)^2, & |t| < a, \\ 0, & |t| > a, \end{cases} \quad \text{is} \quad F(\omega) = \frac{4\cos(\omega a)}{\omega^2 a} - \frac{4\sin(\omega a)}{\omega^3 a^2} + \frac{2\sin(\omega a)}{\omega}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

10. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} 1 - t/\tau, & 0 \le t < 2\tau, \\ 0, & \text{otherwise}, \end{cases} \qquad \text{is} \qquad F(\omega) = \frac{2e^{-i\omega\tau}}{i\omega}\left[\frac{\sin(\omega\tau)}{\omega\tau} - \cos(\omega\tau)\right].$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

11. (a) Show that the Fourier transform of

$$f(t) = \begin{cases} 1 - (t/a)^2, & |t| \le a, \\ 0, & |t| \ge a, \end{cases} \qquad \text{and} \qquad F(\omega) = \frac{4\sin(\omega a) - 4a\omega\cos(\omega a)}{a^2\omega^3}.$$

Using MATLAB, plot the amplitude and phase spectra for this transform.

(b) Rewrite $f(t)$ in terms of step functions. Then use MATLAB's `fourier` to find $F(\omega)$.

12. The integral representation[5] of the modified Bessel function $K_\nu(\cdot)$ is

$$K_\nu\left(a|\omega|\right) = \frac{\Gamma\left(\nu+\frac{1}{2}\right)(2a)^\nu}{|\omega|^\nu \Gamma\left(\frac{1}{2}\right)} \int_0^\infty \frac{\cos(\omega t)}{(t^2+a^2)^{\nu+1/2}}\,dt,$$

[5] Watson, G. N., 1966: *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, p. 185.

where $\Gamma(\cdot)$ is the gamma function, $\nu \geq 0$ and $a > 0$. Use this relationship to show that

$$\mathcal{F}\left[\frac{1}{(t^2+a^2)^{\nu+1/2}}\right] = \frac{2|\omega|^{\nu}\Gamma\left(\frac{1}{2}\right)K_{\nu}\left(a|\omega|\right)}{\Gamma\left(\nu+\frac{1}{2}\right)(2a)^{\nu}}.$$

13. Show that the Fourier transform of a constant K is $2\pi\delta(\omega)K$.

14. Show that

$$\int_a^b \tau\,\delta(t-\tau)\,d\tau = t\left[H(t-a) - H(t-b)\right].$$

Hint: Use integration by parts.

15. For the real function $f(t)$ with Fourier transform $F(\omega)$, prove that $|F(\omega)| = |F(-\omega)|$ and the phase of $F(\omega)$ is an odd function of ω.

11.2 FOURIER TRANSFORMS CONTAINING THE DELTA FUNCTION

In the previous section we stressed the fact that such simple functions as cosine and sine are not absolutely integrable. Does this mean that these functions do not possess a Fourier transform? In this section we shall show that certain functions can still have a Fourier transform even though we cannot compute them directly.

The reason why we can find the Fourier transform of certain functions that are not absolutely integrable lies with the introduction of the delta function because

$$\int_{-\infty}^{\infty} \delta(\omega-\omega_0)e^{it\omega}\,d\omega = e^{i\omega_0 t} \tag{11.2.1}$$

for all t. Thus, the inverse of the Fourier transform $\delta(\omega-\omega_0)$ is the complex exponential $e^{i\omega_0 t}/2\pi$ or

$$\mathcal{F}\left(e^{i\omega_0 t}\right) = 2\pi\delta(\omega-\omega_0). \tag{11.2.2}$$

This immediately yields the result that

$$\mathcal{F}\left(1\right) = 2\pi\delta(\omega), \tag{11.2.3}$$

if we set $\omega_0 = 0$. Thus, the Fourier transform of 1 is an impulse at $\omega = 0$ with weight 2π. Because the Fourier transform equals zero for all $\omega \neq 0$, $f(t) = 1$ does not contain a nonzero frequency and is consequently a DC signal.

Another set of transforms arises from Euler's formula because we have that

$$\mathcal{F}[\sin(\omega_0 t)] = \left[\mathcal{F}\left(e^{i\omega_0 t}\right) - \mathcal{F}\left(e^{-i\omega_0 t}\right)\right]/(2i) \tag{11.2.4}$$

$$= \pi\left[\delta(\omega-\omega_0) - \delta(\omega+\omega_0)\right]/i \tag{11.2.5}$$

$$= -\pi i\delta(\omega-\omega_0) + \pi i\delta(\omega+\omega_0) \tag{11.2.6}$$

and

$$\mathcal{F}[\cos(\omega_0 t)] = \tfrac{1}{2}\left[\mathcal{F}\left(e^{i\omega_0 t}\right) + \mathcal{F}\left(e^{-i\omega_0 t}\right)\right] = \pi\left[\delta(\omega-\omega_0) + \delta(\omega+\omega_0)\right]. \tag{11.2.7}$$

Note that although the amplitude spectra of $\sin(\omega_0 t)$ and $\cos(\omega_0 t)$ are the same, their phase spectra are different.

Let us consider the Fourier transform of any arbitrary periodic function. Recall that any such function $f(t)$ with period $2L$ can be rewritten as the complex Fourier series

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega_0 t}, \tag{11.2.8}$$

where $\omega_0 = \pi/L$. The Fourier transform of $f(t)$ is

$$F(\omega) = \mathcal{F}[f(t)] = \sum_{n=-\infty}^{\infty} 2\pi c_n \delta(\omega - n\omega_0). \tag{11.2.9}$$

Therefore, the Fourier transform of any arbitrary periodic function is a sequence of impulses with weight $2\pi c_n$ located at $\omega = n\omega_0$ with $n = 0, \pm1, \pm2, \ldots$. Thus, the Fourier series and transform of a periodic function are closely related.

• Example 11.2.1: Fourier transform of the sign function

Consider the sign function

$$\operatorname{sgn}(t) = \begin{cases} 1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0. \end{cases} \tag{11.2.10}$$

The function is not absolutely integrable. However, let us approximate it by $e^{-\epsilon|t|}\operatorname{sgn}(t)$, where ϵ is a small positive number. This new function is absolutely integrable and we have that

$$\mathcal{F}[\operatorname{sgn}(t)] = \lim_{\epsilon\to0}\left[-\int_{-\infty}^{0} e^{\epsilon t}e^{-i\omega t}\,dt + \int_0^\infty e^{-\epsilon t}e^{-i\omega t}\,dt\right] = \lim_{\epsilon\to0}\left(\frac{-1}{\epsilon - i\omega} + \frac{1}{\epsilon + i\omega}\right). \tag{11.2.11}$$

If $\omega \neq 0$, Equation 11.2.11 equals $2/i\omega$. If $\omega = 0$, Equation 11.2.11 equals 0 because

$$\lim_{\epsilon\to0}\left(\frac{-1}{\epsilon} + \frac{1}{\epsilon}\right) = 0. \tag{11.2.12}$$

Thus, we conclude that

$$\mathcal{F}[\operatorname{sgn}(t)] = \begin{cases} 2/i\omega, & \omega \neq 0, \\ 0, & \omega = 0. \end{cases} \tag{11.2.13}$$

□

• Example 11.2.2: Fourier transform of the step function

An important function in transform methods is the *(Heaviside) step function*

$$H(t) = \begin{cases} 1, & t > 0, \\ 0, & t < 0. \end{cases} \tag{11.2.14}$$

In terms of the sign function it can be written

$$H(t) = \tfrac{1}{2} + \tfrac{1}{2}\operatorname{sgn}(t). \tag{11.2.15}$$

Because the Fourier transforms of 1 and sgn(t) are $2\pi\delta(\omega)$ and $2/i\omega$, respectively, we have that

$$\mathcal{F}[H(t)] = \pi\delta(\omega) + \frac{1}{i\omega}. \qquad \textbf{(11.2.16)}$$

These transforms are used in engineering but the presence of the delta function requires extra care to ensure their proper use.

Problems

1. Verify that

$$\mathcal{F}[\sin(\omega_0 t)H(t)] = \frac{\omega_0}{\omega_0^2 - \omega^2} + \frac{\pi i}{2}[\delta(\omega + \omega_0) - \delta(\omega - \omega_0)].$$

2. Verify that

$$\mathcal{F}[\cos(\omega_0 t)H(t)] = \frac{i\omega}{\omega_0^2 - \omega^2} + \frac{\pi}{2}[\delta(\omega + \omega_0) + \delta(\omega - \omega_0)].$$

3. Using the definition of Fourier transforms and Equation 11.2.17, show that

$$\int_0^\infty e^{-i\omega t}\,dt = \pi\delta(\omega) - \frac{i}{\omega}, \qquad \text{or} \qquad \int_0^\infty e^{i\omega t}\,dt = \pi\delta(\omega) + \frac{i}{\omega}.$$

4. Following Example 11.2.1, show that

$$\mathcal{F}[\text{sgn}(t)\sin(\omega_0 t)] = \frac{2\omega_0}{\omega_0^2 - \omega^2}, \qquad \text{and} \qquad \mathcal{F}[\text{sgn}(t)\cos(\omega_0 t)] = \frac{2\omega i}{\omega_0^2 - \omega^2}.$$

11.3 PROPERTIES OF FOURIER TRANSFORMS

In principle we can compute any Fourier transform from its definition. However, it is far more efficient to derive some simple relationships that relate transforms to each other. This is the purpose of this section.

Linearity

If $f(t)$ and $g(t)$ are functions with Fourier transforms $F(\omega)$ and $G(\omega)$, respectively, then

$$\mathcal{F}[c_1 f(t) + c_2 g(t)] = c_1 F(\omega) + c_2 G(\omega), \qquad \textbf{(11.3.1)}$$

where c_1 and c_2 are (real or complex) constants.

This result follows from the integral definition

$$\mathcal{F}[c_1 f(t) + c_2 g(t)] = \int_{-\infty}^{\infty} [c_1 f(t) + c_2 g(t)]e^{-i\omega t}dt \qquad \textbf{(11.3.2)}$$

$$= c_1 \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt + c_2 \int_{-\infty}^{\infty} g(t)e^{-i\omega t}dt \qquad \textbf{(11.3.3)}$$

$$= c_1 F(\omega) + c_2 G(\omega). \qquad \textbf{(11.3.4)}$$

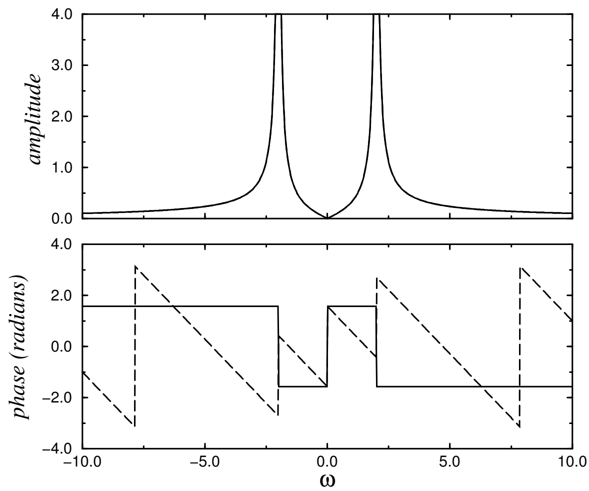


Figure 11.3.1: The amplitude and phase spectra of the Fourier transform for $\cos(2t)\ H(t)$ (solid line) and $\cos[2(t-1)]H(t-1)$ (dashed line). The amplitude becomes infinite at $\omega = \pm 2$.

Time shifting

If $f(t)$ is a function with a Fourier transform $F(\omega)$, then $\mathcal{F}[f(t-\tau)] = e^{-i\omega\tau}F(\omega)$. This follows from the definition of the Fourier transform

$$\mathcal{F}[f(t-\tau)] = \int_{-\infty}^{\infty} f(t-\tau)e^{-i\omega t}\,dt = \int_{-\infty}^{\infty} f(x)e^{-i\omega(x+\tau)}\,dx \tag{11.3.5}$$

$$= e^{-i\omega\tau}\int_{-\infty}^{\infty} f(x)e^{-i\omega x}\,dx = e^{-i\omega\tau}F(\omega). \tag{11.3.6}$$

• Example 11.3.1

The Fourier transform of $f(t) = \cos(at)H(t)$ is $F(\omega) = i\omega/(a^2-\omega^2) + \pi[\delta(\omega+a) + \delta(\omega-a)]/2$. Therefore,

$$\mathcal{F}\{\cos[a(t-k)]H(t-k)\} = e^{-ik\omega}\mathcal{F}[\cos(at)H(t)], \tag{11.3.7}$$

or

$$\mathcal{F}\{\cos[a(t-k)]H(t-k)\} = \frac{i\omega e^{-ik\omega}}{a^2-\omega^2} + \frac{\pi}{2}e^{-ik\omega}[\delta(\omega+a)+\delta(\omega-a)]. \tag{11.3.8}$$

In Figure 11.3.1 we present the amplitude and phase spectra for $\cos(2t)\ H(t)$ (the solid line) while the dashed line gives these spectra for $\cos[2(t-1)]H(t-1)$. This figure shows that the amplitude spectra are identical (why?) while the phase spectra are considerably different. □

Scaling factor

Let $f(t)$ be a function with a Fourier transform $F(\omega)$ and k be a real, nonzero constant. Then $\mathcal{F}[f(kt)] = F(\omega/k)/|k|$.

From the definition of the Fourier transform:

$$\mathcal{F}[f(kt)] = \int_{-\infty}^{\infty} f(kt)e^{-i\omega t}\,dt = \frac{1}{|k|}\int_{-\infty}^{\infty} f(x)e^{-i(\omega/k)x}\,dx = \frac{1}{|k|}F\left(\frac{\omega}{k}\right). \qquad (\mathbf{11.3.9})$$

• **Example 11.3.2**

The Fourier transform of $f(t) = e^{-t}H(t)$ is $F(\omega) = 1/(1+\omega i)$. Therefore, the Fourier transform for $f(at) = e^{-at}H(t)$, $a > 0$, is

$$\mathcal{F}[f(at)] = \left(\frac{1}{a}\right)\left(\frac{1}{1+i\omega/a}\right) = \frac{1}{a+\omega i}. \qquad (\mathbf{11.3.10})$$

To illustrate this scaling property we use the MATLAB script

```
clear; % clear all previous computations
omega = [-10:0.01:10]; % set up frequencies
% real part of transform with a = 1
f1r_omega = 1./(1+omega.*omega);
% imaginary part of transform with a = 1
f1i_omega = - omega./(1+omega.*omega);
% real part of transform with a = 2
f2r_omega = 2./(4+omega.*omega);
% imaginary part of transform with a = 2
f2i_omega = - omega./(4+omega.*omega);
% compute the amplitude of the first transform
ampl1 = sqrt(f1r_omega.*f1r_omega + f1i_omega.*f1i_omega);
% compute the amplitude of the second transform
ampl2 = sqrt(f2r_omega.*f2r_omega + f2i_omega.*f2i_omega);
% compute phase of first transform
phase1 = atan2(f1i_omega,f1r_omega);
% compute phase of second transform
phase2 = atan2(f2i_omega,f2r_omega);
clf; % clear all previous figures
% plot amplitudes of Fourier transforms
subplot(2,1,1), plot(omega,ampl1,omega,ampl2,'--')
ylabel('|F(\omega)|','FontSize',15) % label amplitude plot
% plot phases of Fourier transforms
subplot(2,1,2), plot(omega,phase1,omega,phase2,'--')
ylabel('phase','FontSize',15) % label amplitude plot
xlabel('\omega','FontSize',15) % label x-axis
```

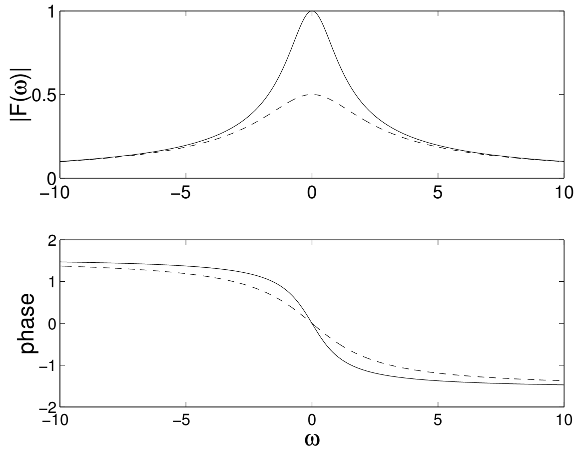


Figure 11.3.2: The amplitude and phase spectra of the Fourier transform for $e^{-t}H(t)$ (solid line) and $e^{-2t}H(t)$ (dashed line).

to plot the amplitude and phase when $a = 1$ and $a = 2$. Figure 11.3.2 shows the results from the MATLAB script: The amplitude spectra decreased by a factor of two for $e^{-2t}H(t)$ compared to $e^{-t}H(t)$ while the differences in the phase are smaller. □

Symmetry

If the function $f(t)$ has the Fourier transform $F(\omega)$, then $\mathcal{F}[F(t)] = 2\pi f(-\omega)$. From the definition of the inverse Fourier transform,

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(x)e^{ixt}\,dx. \qquad \textbf{(11.3.11)}$$

Then

$$2\pi f(-\omega) = \int_{-\infty}^{\infty} F(x)e^{-i\omega x}\,dx = \int_{-\infty}^{\infty} F(t)e^{-i\omega t}\,dt = \mathcal{F}[F(t)]. \qquad \textbf{(11.3.12)}$$

• **Example 11.3.3**

The Fourier transform of $1/(1+t^2)$ is $\pi e^{-|\omega|}$. Therefore,

$$\mathcal{F}\left(\pi e^{-|t|}\right) = \frac{2\pi}{1+\omega^2} \qquad \text{or} \qquad \mathcal{F}\left(e^{-|t|}\right) = \frac{2}{1+\omega^2}. \qquad \textbf{(11.3.13)}$$

□

Derivatives of functions

Let $f^{(k)}(t), k = 0, 1, 2, \ldots, n-1$, be continuous and $f^{(n)}(t)$ be piecewise continuous. Let $|f^{(k)}(t)| \le Ke^{-bt}, b > 0, 0 \le t < \infty; |f^{(k)}(t)| \le Me^{at}, a > 0, -\infty < t \le 0, k = 0, 1, \ldots, n$. Then, $\mathcal{F}[f^{(n)}(t)] = (i\omega)^n F(\omega)$.

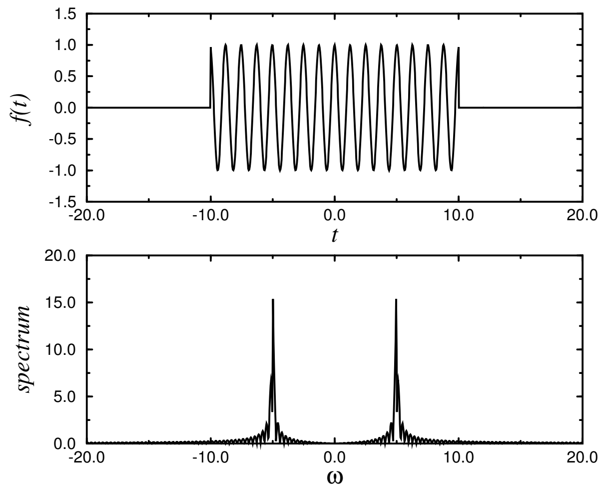


Figure 11.3.3: The (amplitude) spectrum of a rectangular pulse Equation 11.1.9 with a half width $a = 10$ that has been modulated with $\cos(5t)$.

We begin by noting that if the transform $\mathcal{F}[f'(t)]$ exists, then

$$\mathcal{F}[f'(t)] = \int_{-\infty}^{\infty} f'(t)e^{-i\omega t}dt = \int_{-\infty}^{\infty} f'(t)e^{\omega_i t}[\cos(\omega_r t) - i\sin(\omega_r t)]\,dt \tag{11.3.14}$$

$$= (-\omega_i + i\omega_r)\int_{-\infty}^{\infty} f(t)e^{\omega_i t}[\cos(\omega_r t) - i\sin(\omega_r t)]\,dt \tag{11.3.15}$$

$$= i\omega\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt = i\omega F(\omega). \tag{11.3.16}$$

Finally,

$$\mathcal{F}[f^{(n)}(t)] = i\omega\mathcal{F}[f^{(n-1)}(t)] = (i\omega)^2\mathcal{F}[f^{(n-2)}(t)] = \cdots = (i\omega)^n F(\omega). \tag{11.3.17}$$

• Example 11.3.4

The Fourier transform of $f(t) = 1/(1+t^2)$ is $F(\omega) = \pi e^{-|\omega|}$. Therefore,

$$\mathcal{F}\left[-\frac{2t}{(1+t^2)^2}\right] = i\omega\pi e^{-|\omega|}, \qquad \text{or} \qquad \mathcal{F}\left[\frac{t}{(1+t^2)^2}\right] = -\frac{i\omega\pi}{2}e^{-|\omega|}. \tag{11.3.18}$$

□

Modulation

In communications a popular method of transmitting information is by *amplitude modulation* (AM). In this process the signal is carried according to the expression $f(t)e^{i\omega_0 t}$, where ω_0 is the *carrier frequency* and $f(t)$ is an arbitrary function of time whose amplitude spectrum peaks at some frequency that is usually small compared to ω_0. We now show that

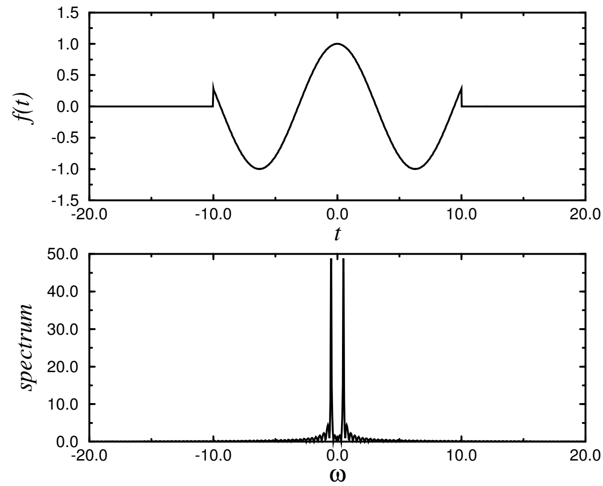


Figure 11.3.4: The (amplitude) spectrum of a rectangular pulse Equation 11.1.9 with a half width $a = 10$ that has been modulated with $\cos(t/2)$.

the Fourier transform of $f(t)e^{i\omega_0 t}$ is $F(\omega - \omega_0)$, where $F(\omega)$ is the Fourier transform of $f(t)$.

We begin by using the definition of the Fourier transform, or

$$\mathcal{F}[f(t)e^{i\omega_0 t}] = \int_{-\infty}^{\infty} f(t)e^{i\omega_0 t}e^{-i\omega t}dt = \int_{-\infty}^{\infty} f(t)e^{-i(\omega-\omega_0)t}dt = F(\omega - \omega_0). \qquad \textbf{(11.3.19)}$$

Therefore, if we have the spectrum of a particular function $f(t)$, then the Fourier transform of the modulated function $f(t)e^{i\omega_0 t}$ is the same as that for $f(t)$ except that it is now centered on the frequency ω_0 rather than on the zero frequency.

• Example 11.3.5

Let us determine the Fourier transform of a square pulse modulated by a cosine wave as shown in Figures 11.3.3 and 11.3.4. Because $\cos(\omega_0 t) = \frac{1}{2}[e^{i\omega_0 t} + e^{-i\omega_0 t}]$ and the Fourier transform of a square pulse is $F(\omega) = 2\sin(\omega a)/\omega$,

$$\mathcal{F}[f(t)\cos(\omega_0 t)] = \frac{\sin[(\omega - \omega_0)a]}{\omega - \omega_0} + \frac{\sin[(\omega + \omega_0)a]}{\omega + \omega_0}. \qquad \textbf{(11.3.20)}$$

Therefore, the Fourier transform of the modulated pulse equals one half of the sum of the Fourier transform of the pulse centered on ω_0 and $-\omega_0$. See Figures 11.3.3 and 11.3.4.

In many practical situations, $\omega_0 \gg \pi/a$. In this case we may treat each term as completely independent from the other; the contribution from the peak at $\omega = \omega_0$ has a negligible effect on the peak at $\omega = -\omega_0$. □

• Example 11.3.6

The Fourier transform of $f(t) = e^{-bt}H(t)$ is $F(\omega) = 1/(b + i\omega)$. Therefore,

$$\mathcal{F}[e^{-bt}\cos(at)H(t)] = \tfrac{1}{2}\mathcal{F}\left(e^{iat}e^{-bt} + e^{-iat}e^{-bt}\right) \qquad \textbf{(11.3.21)}$$

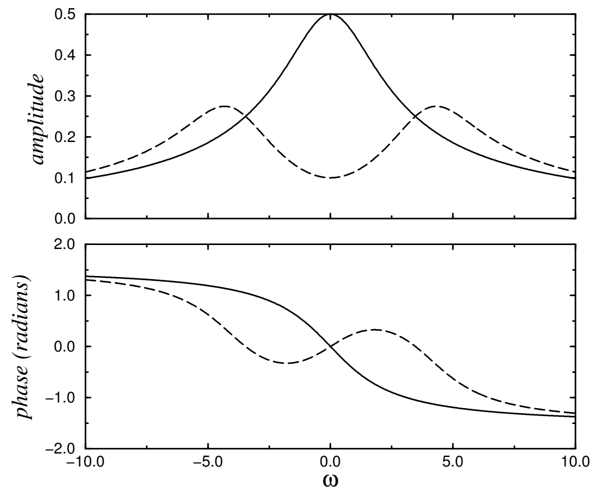


Figure 11.3.5: The amplitude and phase spectra of the Fourier transform for $e^{-2t}\,H(t)$ (solid line) and $e^{-2t}\cos(4t)H(t)$ (dashed line).

$$\mathcal{F}[e^{-bt}\cos(at)H(t)] = \frac{1}{2}\left(\left.\frac{1}{b+i\omega'}\right|_{\omega'=\omega-a} + \left.\frac{1}{b+i\omega'}\right|_{\omega'=\omega+a}\right) \tag{11.3.22}$$

$$= \frac{1}{2}\left[\frac{1}{(b+i\omega)-ai} + \frac{1}{(b+i\omega)+ai}\right] \tag{11.3.23}$$

$$= \frac{b+i\omega}{(b+i\omega)^2+a^2}. \tag{11.3.24}$$

We illustrate this result using $e^{-2t}H(t)$ and $e^{-2t}\cos(4t)H(t)$ in Figure 11.3.5.

□

• Example 11.3.7: Frequency modulation

In contrast to amplitude modulation, *frequency modulation* (FM) transmits information by instantaneous variations of the carrier frequency. It can be expressed mathematically as $\exp\left[i\int_{-\infty}^t f(\tau)\,d\tau + iC\right] e^{i\omega_0 t}$, where C is a constant. To illustrate this concept, let us find the Fourier transform of a simple frequency modulation

$$f(t) = \begin{cases} \omega_1, & |t| < T/2, \\ 0, & |t| > T/2, \end{cases} \tag{11.3.25}$$

and $C = -\omega_1 T/2$. In this case, the signal in the time domain is

$$g(t) = \exp\left[i\int_{-\infty}^t f(\tau)\,d\tau + iC\right] e^{i\omega_0 t} = \begin{cases} e^{-i\omega_1 T/2}e^{i\omega_0 t}, & t < -T/2, \\ e^{i\omega_1 t}e^{i\omega_0 t}, & -T/2 < t < T/2, \\ e^{i\omega_1 T/2}e^{i\omega_0 t}, & T/2 < t. \end{cases} \tag{11.3.26}$$

We illustrate this signal in Figures 11.3.6 and 11.3.7.

The Fourier transform of the signal $G(\omega)$ equals

$$G(\omega) = e^{-i\omega_1 T/2}\int_{-\infty}^{-T/2} e^{i(\omega_0-\omega)t}\,dt + \int_{-T/2}^{T/2} e^{i(\omega_0+\omega_1-\omega)t}\,dt + e^{i\omega_1 T/2}\int_{T/2}^{\infty} e^{i(\omega_0-\omega)t}\,dt$$

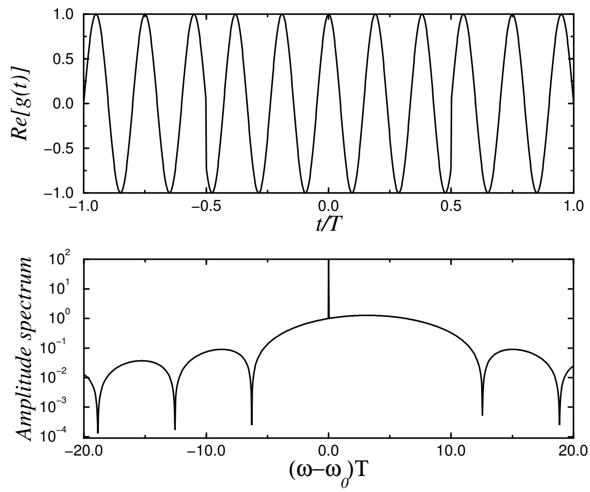


Figure 11.3.6: The (amplitude) spectrum $|G(\omega)|/T$ of a frequency-modulated signal (shown top) when $\omega_1 T = 2\pi$ and $\omega_0 T = 10\pi$. The transform becomes undefined at $\omega = \omega_0$.

$$= e^{-i\omega_1 T/2}\int_{-\infty}^{0} e^{i(\omega_0-\omega)t}\,dt + e^{i\omega_1 T/2}\int_{0}^{\infty} e^{i(\omega_0-\omega)t}\,dt - e^{-i\omega_1 T/2}\int_{-T/2}^{0} e^{i(\omega_0-\omega)t}\,dt \tag{11.3.27}$$

$$+ \int_{-T/2}^{T/2} e^{i(\omega_0+\omega_1-\omega)t}\,dt - e^{i\omega_1 T/2}\int_{0}^{T/2} e^{i(\omega_0-\omega)t}\,dt. \tag{11.3.28}$$

Applying the fact that

$$\int_0^\infty e^{\pm i\alpha t}\,dt = \pi\delta(\alpha) \pm \frac{i}{\alpha}, \tag{11.3.29}$$

$$G(\omega) = \pi\delta(\omega-\omega_0)\left[e^{i\omega_1 T/2} + e^{-i\omega_1 T/2}\right] + \frac{\left[e^{i(\omega_0+\omega_1-\omega)T/2} - e^{-i(\omega_0+\omega_1-\omega)T/2}\right]}{i(\omega_0+\omega_1-\omega)}$$
$$- \frac{\left[e^{i(\omega_0+\omega_1-\omega)T/2} - e^{-i(\omega_0+\omega_1-\omega)T/2}\right]}{i(\omega_0-\omega)} \tag{11.3.30}$$
$$= 2\pi\delta(\omega-\omega_0)\cos(\omega_1 T/2) + \frac{2\omega_1 \sin[(\omega-\omega_0-\omega_1)T/2]}{(\omega-\omega_0)(\omega-\omega_0-\omega_1)}. \tag{11.3.31}$$

Figures 11.3.6 and 11.3.7 illustrate the amplitude spectrum for various parameters. In general, the transform is not symmetric, with an increasing number of humped curves as $\omega_1 T$ increases. □

Parseval's equality

In applying Fourier methods to practical problems we may encounter a situation where we are interested in computing the energy of a system. Energy is usually expressed by the integral $\int_{-\infty}^{\infty} |f(t)|^2\,dt$. Can we compute this integral if we only have the Fourier transform of $F(\omega)$?

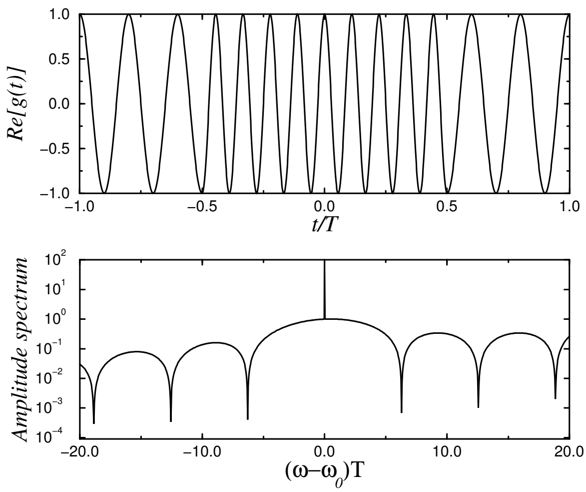


Figure 11.3.7: The (amplitude) spectrum $|G(\omega)|/T$ of a frequency-modulated signal (shown top) when $\omega_1 T = 8\pi$ and $\omega_0 T = 10\pi$. The transform becomes undefined at $\omega = \omega_0$.

From the definition of the inverse Fourier transform

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega, \qquad \textbf{(11.3.32)}$$

we have that

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)\left[\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega\right] dt. \qquad \textbf{(11.3.33)}$$

Interchanging the order of integration on the right side of Equation 11.3.33,

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\left[\int_{-\infty}^{\infty} f(t)e^{i\omega t}dt\right] d\omega. \qquad \textbf{(11.3.34)}$$

However,

$$F^*(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt. \qquad \textbf{(11.3.35)}$$

Therefore,

$$\boxed{\int_{-\infty}^{\infty} |f(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2\, d\omega.} \qquad \textbf{(11.3.36)}$$

This is *Parseval's equality*[6] as it applies to Fourier transforms. The quantity $|F(\omega)|^2$ is called the *power spectrum*.

[6] Apparently first derived by Rayleigh, J. W., 1889: On the character of the complete radiation at a given temperature. *Philos. Mag., Ser. 5*, **27**, 460–469.

Some General Properties of Fourier Transforms

	function, f(t)	**Fourier transform, F(ω)**
1. Linearity	$c_1f(t) + c_2g(t)$	$c_1F(\omega) + c_2G(\omega)$
2. Complex conjugate	$f^*(t)$	$F^*(-\omega)$
3. Scaling	$f(\alpha t)$	$F(\omega/\alpha)/\lvert\alpha\rvert$
4. Delay	$f(t-\tau)$	$e^{-i\omega\tau}F(\omega)$
5. Frequency translation	$e^{i\omega_0 t}f(t)$	$F(\omega-\omega_0)$
6. Duality-time frequency	$F(t)$	$2\pi f(-\omega)$
7. Time differentiation	$f'(t)$	$i\omega F(\omega)$

• **Example 11.3.8**

In Example 11.1.1, we showed that the Fourier transform for a unit rectangular pulse between $-a < t < a$ is $2\sin(\omega a)/\omega$. Therefore, by Parseval's equality,

$$\frac{2}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2(\omega a)}{\omega^2}\,d\omega = \int_{-a}^{a} 1^2\,dt = 2a, \qquad \text{or} \qquad \int_{-\infty}^{\infty}\frac{\sin^2(\omega a)}{\omega^2}\,d\omega = \pi a. \tag{11.3.37}$$

□

Poisson's summation formula

If $f(x)$ is integrable over $(-\infty,\infty)$, there exists a relationship between the function and its Fourier transform, commonly called *Poisson's summation formula*.[7]

We begin by inventing a periodic function $g(x)$ defined by

$$g(x) = \sum_{k=-\infty}^{\infty} f(x+2\pi k). \tag{11.3.38}$$

Because $g(x)$ is a periodic function of 2π, it can be represented by the complex Fourier series:

$$g(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \qquad \text{or} \qquad g(0) = \sum_{k=-\infty}^{\infty} f(2\pi k) = \sum_{n=-\infty}^{\infty} c_n. \tag{11.3.39}$$

[7] Poisson, S. D., 1823: Suite du mémoire sur les intégrales définies et sur la sommation des séries. *J. École Polytech.*, **19**, 404–509. See page 451.

Computing c_n, we find that

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)e^{-inx}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=-\infty}^{\infty} f(x+2k\pi)e^{-inx}\,dx \tag{11.3.40}$$

$$= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\int_{-\pi}^{\pi} f(x+2k\pi)e^{-inx}\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-inx}\,dx = \frac{F(n)}{2\pi}, \tag{11.3.41}$$

where $F(\omega)$ is the Fourier transform of $f(x)$. Substituting Equation 11.3.41 into the right side of Equation 11.3.39, we obtain

$$\sum_{k=-\infty}^{\infty} f(2\pi k) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} F(n) \tag{11.3.42}$$

or

$$\sum_{k=-\infty}^{\infty} f(\alpha k) = \frac{1}{\alpha}\sum_{n=-\infty}^{\infty} F\left(\frac{2\pi n}{\alpha}\right). \tag{11.3.43}$$

• **Example 11.3.9**

One of the popular uses of Poisson's summation formula is the evaluation of infinite series. For example, let $f(x) = 1/(a^2+x^2)$ with a real and nonzero. Then, $F(\omega) = \pi e^{-|a\omega|}/|a|$ and

$$\sum_{k=-\infty}^{\infty} \frac{1}{a^2+(2\pi k)^2} = \frac{1}{2}\sum_{n=-\infty}^{\infty}\frac{1}{|a|}e^{-|an|} = \frac{1}{2|a|}\left(1+2\sum_{n=1}^{\infty}e^{-|a|n}\right) \tag{11.3.44}$$

$$= \frac{1}{2|a|}\left(-1+\frac{2}{1-e^{-|a|}}\right) = \frac{1}{2|a|}\coth\left(\frac{|a|}{2}\right). \tag{11.3.45}$$

Problems

1. Find the Fourier transform of $1/(1+a^2t^2)$, where a is real, given that $\mathcal{F}[1/(1+t^2)] = \pi e^{-|\omega|}$.

2. Find the Fourier transform of $\cos(at)/(1+t^2)$, where a is real, given that $\mathcal{F}[1/(1+t^2)] = \pi e^{-|\omega|}$.

3. Use the fact that $\mathcal{F}[e^{-at}H(t)] = 1/(a+i\omega)$ with $a>0$ and Parseval's equality to show that

$$\int_{-\infty}^{\infty}\frac{dx}{x^2+a^2} = \frac{\pi}{a}.$$

4. Use the fact that $\mathcal{F}[1/(1+t^2)] = \pi e^{-|\omega|}$ and Parseval's equality to show that

$$\int_{-\infty}^{\infty} \frac{dx}{(x^2+1)^2} = \frac{\pi}{2}.$$

5. Use the function $f(t) = e^{-at}\sin(bt)H(t)$ with $a > 0$ and Parseval's equality to show that

$$2\int_0^{\infty} \frac{dx}{(x^2+a^2-b^2)^2+4a^2b^2} = \int_{-\infty}^{\infty} \frac{dx}{(x^2+a^2-b^2)^2+4a^2b^2} = \frac{\pi}{2a(a^2+b^2)}.$$

6. Using the modulation property and $\mathcal{F}[e^{-bt}H(t)] = 1/(b+i\omega)$, show that

$$\mathcal{F}\left[e^{-bt}\sin(at)H(t)\right] = \frac{a}{(b+i\omega)^2+a^2}.$$

Use MATLAB to plot and compare the amplitude and phase spectra for $e^{-t}\ H(t)$ and $e^{-t}\sin(2t)\ H(t)$.

7. Use Poisson's summation formula with $f(t) = e^{-|t|}$ to show that

$$\sum_{n=-\infty}^{\infty} \frac{1}{n^2+1} = \pi\frac{1+e^{-2\pi}}{1-e^{-2\pi}}.$$

8. Use Poisson's summation formula to prove[8] that

$$\sum_{n=-\infty}^{\infty} e^{-a(n+c)^2+2b(n+c)} = \sqrt{\frac{\pi}{a}}e^{b^2/a}\sum_{n=-\infty}^{\infty} e^{-n^2\pi^2/a-2n\pi i(b/a-c)}.$$

9. Use Poisson's summation formula to prove that

$$\sum_{n=-\infty}^{\infty} e^{-ianT} = \frac{2\pi}{T}\sum_{n=-\infty}^{\infty} \delta\left(\frac{2\pi n}{T}-a\right),$$

where $\delta(\cdot)$ is the Dirac delta function.

10. Prove the two-dimensional form[9] of Poisson's summation formula:

$$\sum_{k_1=-\infty}^{\infty}\sum_{k_2=-\infty}^{\infty} f(\alpha_1 k_1, \alpha_2 k_2) = \frac{1}{\alpha_1\alpha_2}\sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty} F\left(\frac{2\pi n_1}{\alpha_1}, \frac{2\pi n_2}{\alpha_2}\right),$$

[8] First proved by Ewald, P. P., 1921: Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys., 4te Folge*, **64**, 253–287.

[9] Lucas, S. K., R. Sipcic, and H. A. Stone, 1997: An integral equation solution for the steady-state current at a periodic array of surface microelectrodes. *SIAM J. Appl. Math.*, **57**, 1615–1638.

where

$$F(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-i\omega_1 x - i\omega_2 y}\, dx\, dy.$$

11.4 INVERSION OF FOURIER TRANSFORMS

Having focused on the Fourier transform in the previous sections, we now consider the inverse Fourier transform. Recall that the improper integral, Equation 11.1.6, defines the inverse. Consequently, one method of inversion is direct integration.

• Example 11.4.1

Let us find the inverse of $F(\omega) = \pi e^{-|\omega|}$.
From the definition of the inverse Fourier transform,

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \pi e^{-|\omega|} e^{i\omega t} d\omega = \frac{1}{2} \int_{-\infty}^{0} e^{(1+it)\omega} d\omega + \frac{1}{2} \int_{0}^{\infty} e^{(-1+it)\omega} d\omega \qquad \textbf{(11.4.1)}$$

$$= \frac{1}{2} \left[\frac{e^{(1+it)\omega}}{1+it} \bigg|_{-\infty}^{0} + \frac{e^{(-1+it)\omega}}{-1+it} \bigg|_{0}^{\infty} \right] = \frac{1}{2} \left[\frac{1}{1+it} - \frac{1}{-1+it} \right] = \frac{1}{1+t^2}. \qquad \textbf{(11.4.2)}$$

An alternative to direct integration is the MATLAB function `ifourier`. For example, to invert $F(\omega) = \pi e^{-|\omega|}$, we type in the commands:

```
>> syms pi omega t
>> ifourier('pi*exp(-abs(omega))',omega,t)
```

This yields

```
ans =
1/(1+t^2)
```

□

Another method for inverting Fourier transforms is rewriting the Fourier transform using partial fractions so that we can use transform tables. The following example illustrates this technique.

• Example 11.4.2

Let us invert the transform

$$F(\omega) = \frac{1}{(1+i\omega)(1-2i\omega)^2}. \qquad \textbf{(11.4.3)}$$

We begin by rewriting (11.4.3) as

$$F(\omega) = \frac{1}{9} \left[\frac{1}{1+i\omega} + \frac{2}{1-2i\omega} + \frac{6}{(1-2i\omega)^2} \right] = \frac{1}{9(1+i\omega)} + \frac{1}{9(\frac{1}{2} - i\omega)} + \frac{1}{6(\frac{1}{2} - i\omega)^2}. \qquad \textbf{(11.4.4)}$$

Using a table of Fourier transforms (see Section 11.1), we invert Equation 11.4.4 term by term and find that

$$f(t) = \tfrac{1}{9} e^{-t} H(t) + \tfrac{1}{9} e^{t/2} H(-t) - \tfrac{1}{6} t e^{t/2} H(-t). \qquad \textbf{(11.4.5)}$$

To check our answer, we type the following commands into MATLAB:

```
>> syms omega t
>> ifourier(1/((1+i*omega)*(1-2*i*omega)^2),omega,t)
```

which yields

```
ans =
1/9*exp(-t)*Heaviside(t)-1/6*exp(1/2*t)*t*Heaviside(-t)
    +1/9*exp(1/2*t)*Heaviside(-t)
```

□

Although we may find the inverse by direct integration or partial fractions, in many instances the Fourier transform does not lend itself to these techniques. On the other hand, if we view the inverse Fourier transform as a line integral along the real axis in the complex ω-plane, then some of the techniques that we developed in Chapter 10 can be applied to this problem. To this end, we rewrite the inversion integral, Equation 11.1.6, as

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{it\omega}\,d\omega = \frac{1}{2\pi}\oint_C F(z)e^{itz}\,dz - \frac{1}{2\pi}\int_{C_R} F(z)e^{itz}\,dz, \qquad \textbf{(11.4.6)}$$

where C denotes a closed contour consisting of the entire real axis plus a new contour C_R that joins the point $(\infty, 0)$ to $(-\infty, 0)$. There are countless possibilities for C_R. For example, it could be the loop $(\infty, 0)$ to (∞, R) to $(-\infty, R)$ to $(-\infty, 0)$ with $R > 0$. However, any choice of C_R must be such that we can compute $\int_{C_R} F(z)e^{itz}\,dz$. When we take that constraint into account, the number of acceptable contours decreases to just a few. The best is given by *Jordan's lemma*.[10]

Jordan's lemma: *Suppose that, on a circular arc C_R with radius R and center at the origin, $f(z) \to 0$ uniformly as $R \to \infty$. Then*

$$(1) \qquad \lim_{R\to\infty}\int_{C_R} f(z)e^{imz}\,dz = 0, \qquad (m > 0) \qquad \textbf{(11.4.7)}$$

if C_R lies in the first and/or second quadrant;

$$(2) \qquad \lim_{R\to\infty}\int_{C_R} f(z)e^{-imz}\,dz = 0, \qquad (m > 0) \qquad \textbf{(11.4.8)}$$

if C_R lies in the third and/or fourth quadrant;

$$(3) \qquad \lim_{R\to\infty}\int_{C_R} f(z)e^{mz}\,dz = 0, \qquad (m > 0) \qquad \textbf{(11.4.9)}$$

if C_R lies in the second and/or third quadrant; and

$$(4) \qquad \lim_{R\to\infty}\int_{C_R} f(z)e^{-mz}\,dz = 0, \qquad (m > 0) \qquad \textbf{(11.4.10)}$$

if C_R lies in the first and/or fourth quadrant.

---

[10] Jordan, C., 1894: *Cours D'Analyse de l'École Polytechnique. Vol. 2*. Gauthier-Villars, pp. 285–286. See also Whittaker, E. T., and G. N. Watson, 1963: *A Course of Modern Analysis*. Cambridge University Press, p. 115.

Technically, only (1) is actually Jordan's lemma while the remaining points are variations.

Proof: We shall prove the first part; the remaining portions follow by analog. We begin by noting that

$$|I_R| = \left|\int_{C_R} f(z)e^{imz}\,dz\right| \le \int_{C_R} |f(z)|\left|e^{imz}\right||dz|. \tag{11.4.11}$$

Now

$$|dz| = R\,d\theta, \quad |f(z)| \le M_R, \tag{11.4.12}$$

$$\left|e^{imz}\right| = \left|\exp(imRe^{\theta i})\right| = |\exp\{imR[\cos(\theta) + i\sin(\theta)]\}| = e^{-mR\sin(\theta)}. \tag{11.4.13}$$

Therefore,

$$|I_R| \le RM_R \int_{\theta_0}^{\theta_1} \exp[-mR\sin(\theta)]\,d\theta, \tag{11.4.14}$$

where $0 \le \theta_0 < \theta_1 \le \pi$. Because the integrand is positive, the right side of Equation 11.4.14 is largest if we take $\theta_0 = 0$ and $\theta_1 = \pi$. Then

$$|I_R| \le RM_R \int_0^{\pi} e^{-mR\sin(\theta)}\,d\theta = 2RM_R \int_0^{\pi/2} e^{-mR\sin(\theta)}\,d\theta. \tag{11.4.15}$$

We cannot evaluate the integrals in Equation 11.4.15 as they stand. However, because $\sin(\theta) \ge 2\theta/\pi$ if $0 \le \theta \le \pi/2$, we can bound the value of the integral by

$$|I_R| \le 2RM_R \int_0^{\pi/2} e^{-2mR\theta/\pi}\,d\theta = \frac{\pi}{m}M_R\left(1 - e^{-mR}\right). \tag{11.4.16}$$

If $m > 0$, $|I_R|$ tends to zero with M_R as $R \to \infty$. □

Consider now the following inversions of Fourier transforms:

• **Example 11.4.3**

For our first example we find the inverse for

$$F(\omega) = \frac{1}{\omega^2 - 2ib\omega - a^2 - b^2}, \qquad a, b > 0. \tag{11.4.17}$$

From the inversion integral,

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{it\omega}}{\omega^2 - 2ib\omega - a^2 - b^2}\,d\omega, \tag{11.4.18}$$

or

$$f(t) = \frac{1}{2\pi}\oint_C \frac{e^{itz}}{z^2 - 2ibz - a^2 - b^2}\,dz - \frac{1}{2\pi}\int_{C_R} \frac{e^{itz}}{z^2 - 2ibz - a^2 - b^2}\,dz, \tag{11.4.19}$$

where C denotes a closed contour consisting of the entire real axis plus C_R. Because $f(z) = 1/(z^2 - 2ibz - a^2 - b^2)$ tends to zero uniformly as $|z| \to \infty$ and $m = t$, the second integral in Equation 11.4.19 vanishes by Jordan's lemma if C_R is a semicircle of infinite

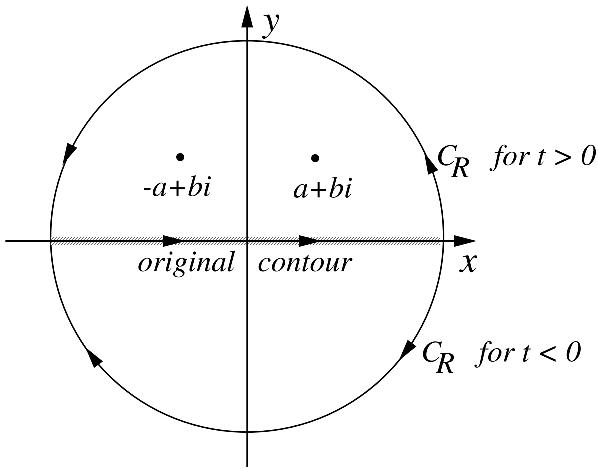


Figure 11.4.1: Contour used to find the inverse of the Fourier transform, Equation 11.4.17. The contour C consists of the line integral along the real axis plus C_R.

radius in the upper half of the z-plane when $t > 0$ and a semicircle in the lower half of the z-plane when $t < 0$.

Next we must find the location and nature of the singularities. They are located at

$$z^2 - 2ibz - a^2 - b^2 = 0, \qquad \text{or} \qquad z = \pm a + bi. \tag{11.4.20}$$

Therefore we can rewrite Equation 11.4.19 as

$$f(t) = \frac{1}{2\pi} \oint_C \frac{e^{itz}}{(z-a-bi)(z+a-bi)}\, dz. \tag{11.4.21}$$

Thus, all of the singularities are simple poles.

Consider now $t > 0$. As stated earlier, we close the line integral with an infinite semicircle in the upper half-plane. See Figure 11.4.1. Inside this closed contour there are two singularities: $z = \pm a + bi$. For these poles,

$$\operatorname{Res}\left(\frac{e^{itz}}{z^2 - 2ibz - a^2 - b^2}; a + bi\right) = \lim_{z \to a+bi} \frac{(z-a-bi)e^{itz}}{(z-a-bi)(z+a-bi)} \tag{11.4.22}$$

$$= \frac{e^{iat}e^{-bt}}{2a} = \frac{e^{-bt}}{2a}[\cos(at) + i\sin(at)], \tag{11.4.23}$$

where we used Euler's formula to eliminate e^{iat}. Similarly,

$$\operatorname{Res}\left(\frac{e^{itz}}{z^2 - 2ibz - a^2 - b^2}; -a + bi\right) = -\frac{e^{-bt}}{2a}[\cos(at) - i\sin(at)]. \tag{11.4.24}$$

Consequently, the inverse Fourier transform follows from Equation 11.4.21 after applying the residue theorem, and equals

$$f(t) = -\frac{e^{-bt}}{2a}\sin(at) \tag{11.4.25}$$

for $t > 0$.

For $t < 0$, the semicircle is in the lower half-plane because the contribution from the semicircle vanishes as $R \to \infty$. Because there are no singularities within the closed contour, $f(t) = 0$. Therefore, we can write in general that

$$f(t) = -\frac{e^{-bt}}{2a} \sin(at) H(t). \tag{11.4.26}$$

□

• Example 11.4.4

Let us find the inverse of the Fourier transform

$$F(\omega) = \frac{e^{-\omega i}}{\omega^2 + a^2}, \tag{11.4.27}$$

where a is real and positive.

From the inversion integral,

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i(t-1)\omega}}{\omega^2 + a^2} \, d\omega = \frac{1}{2\pi} \oint_C \frac{e^{i(t-1)z}}{z^2 + a^2} \, dz - \frac{1}{2\pi} \int_{C_R} \frac{e^{i(t-1)z}}{z^2 + a^2} \, dz, \tag{11.4.28}$$

where C denotes a closed contour consisting of the entire real axis plus C_R. The contour C_R is determined by Jordan's lemma because $1/(z^2 + a^2) \to 0$ uniformly as $|z| \to \infty$. Since $m = t - 1$, the semicircle C_R of infinite radius lies in the upper half-plane if $t > 1$ and in the lower half-plane if $t < 1$. Thus, if $t > 1$,

$$f(t) = \frac{1}{2\pi}(2\pi i) \operatorname{Res}\left[\frac{e^{i(t-1)z}}{z^2 + a^2}; ai\right] = \frac{e^{-a(t-1)}}{2a}, \tag{11.4.29}$$

whereas for $t < 1$,

$$f(t) = \frac{1}{2\pi}(-2\pi i) \operatorname{Res}\left[\frac{e^{i(t-1)z}}{z^2 + a^2}; -ai\right] = \frac{e^{a(t-1)}}{2a}. \tag{11.4.30}$$

The minus sign in front of the $2\pi i$ arises from the clockwise direction or negative sense of the contour. We can write the inverse as the single expression

$$f(t) = \frac{e^{-a|t-1|}}{2a}. \tag{11.4.31}$$

□

• Example 11.4.5

Let us evaluate the integral

$$\int_{-\infty}^{\infty} \frac{\cos(kx)}{x^2 + a^2} \, dx, \tag{11.4.32}$$

where $a, k > 0$.

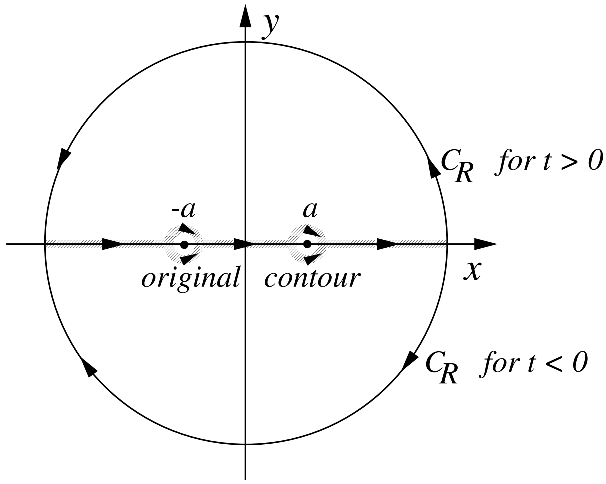


Figure 11.4.2: Contour used in Example 11.4.6.

We begin by noting that

$$\int_{-\infty}^{\infty} \frac{\cos(kx)}{x^2+a^2}\,dx = \Re\left(\int_{-\infty}^{\infty} \frac{e^{ikx}}{x^2+a^2}\,dx\right) = \Re\left(\int_{C_1} \frac{e^{ikz}}{z^2+a^2}\,dz\right), \tag{11.4.33}$$

where C_1 denotes a line integral along the real axis from $-\infty$ to ∞. A quick check shows that the integrand of the right side of Equation 11.4.33 satisfies Jordan's lemma. Therefore,

$$\int_{-\infty}^{\infty} \frac{e^{ikx}}{x^2+a^2}\,dx = \oint_C \frac{e^{ikz}}{z^2+a^2}\,dz = 2\pi i \operatorname{Res}\left(\frac{e^{ikz}}{z^2+a^2}; ai\right) \tag{11.4.34}$$

$$= 2\pi i \lim_{z\to ai} \frac{(z-ai)e^{ikz}}{z^2+a^2} = \frac{\pi}{a}e^{-ka}, \tag{11.4.35}$$

where C denotes the closed infinite semicircle in the upper half-plane. Taking the real and imaginary parts of Equation 11.4.35,

$$\int_{-\infty}^{\infty} \frac{\cos(kx)}{x^2+a^2}\,dx = \frac{\pi}{a}e^{-ka} \qquad \text{and} \qquad \int_{-\infty}^{\infty} \frac{\sin(kx)}{x^2+a^2}\,dx = 0. \tag{11.4.36}$$

□

• Example 11.4.6

Let us now invert the Fourier transform $F(\omega) = 2a/(a^2-\omega^2)$, where a is real. The interesting aspect of this problem is the presence of singularities at $\omega = \pm a$ that lie *along* the contour of integration. How do we use contour integration to compute

$$f(t) = \frac{a}{\pi}\int_{-\infty}^{\infty} \frac{e^{it\omega}}{a^2-\omega^2}\,d\omega? \tag{11.4.37}$$

The answer to this question involves the concept of Cauchy principal value integrals, which allows us to extend the conventional definition of integrals to include integrands that become infinite at a finite number of points. See Section 10.10. Thus, by treating Equation 11.4.37 as a Cauchy principal value integral, we again convert it into a closed contour integration by closing the line integration along the real axis as shown in Figure 11.4.2. The semicircles at infinity vanish by Jordan's lemma and

$$f(t) = \frac{a}{\pi} \oint_C \frac{e^{itz}}{a^2 - z^2}\, dz. \tag{11.4.38}$$

For $t > 0$,

$$f(t) = -\frac{2\pi i a}{\pi}\frac{1}{2}\text{Res}\left[\frac{e^{itz}}{z^2 - a^2}; -a\right] - \frac{2\pi i a}{\pi}\frac{1}{2}\text{Res}\left[\frac{e^{itz}}{z^2 - a^2}; a\right]. \tag{11.4.39}$$

We have the factor $\frac{1}{2}$ because we are only passing over the "top" of the singularity at $z = a$ and $z = -a$. Computing the residues and simplifying the results, we obtain $f(t) = \sin(at)$.

Similarly, when $t < 0$,

$$f(t) = \frac{2\pi i a}{\pi}\frac{1}{2}\text{Res}\left[\frac{e^{itz}}{z^2 - a^2}; -a\right] + \frac{2\pi i a}{\pi}\frac{1}{2}\text{Res}\left[\frac{e^{itz}}{z^2 - a^2}; a\right] = -\sin(at). \tag{11.4.40}$$

These results can be collapsed down to the single expression $f(t) = \text{sgn}(t)\sin(at)$. □

• Example 11.4.7

An additional benefit of understanding inversion by the residue method is the ability to qualitatively anticipate the inverse by knowing the location of the poles of $F(\omega)$. This intuition is important because many engineering analyses discuss stability and performance entirely in terms of the properties of the system's Fourier transform. In Figure 11.4.3 we graphed the location of the poles of $F(\omega)$ and the corresponding $f(t)$. The student should go through the mental exercise of connecting the two pictures.

• Example 11.4.8

So far, we used only the first two points of Jordan's lemma. In this example[11] we illustrate how the remaining two points may be applied.

Consider the contour integral

$$\oint_C \cot(\pi z)\left[\frac{e^{-cz}}{(\tau + 2\pi z)^2 + \beta^2} + \frac{e^{-cz}}{(\tau - 2\pi z)^2 + \beta^2}\right] dz,$$

where $c > 0$ and β, τ are real. Let us evaluate this contour integral where the contour is shown in Figure 11.4.3.

[11] See Hsieh, T. C., and R. Greif, 1972: Theoretical determination of the absorption coefficient and the total band absorptance including a specific application to carbon monoxide. *Int. J. Heat Mass Transfer*, **15**, 1477–1487.

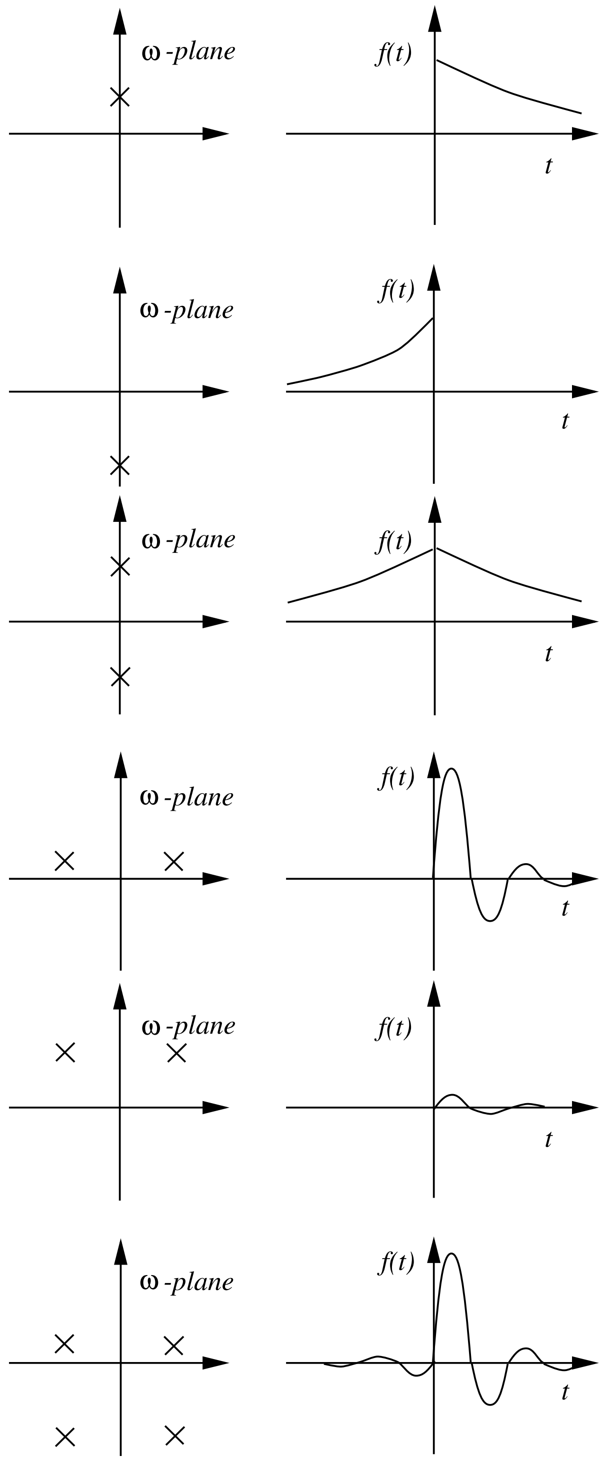


Figure 11.4.3: The correspondence between the location of the simple poles of the Fourier transform $F(\omega)$ and the behavior of $f(t)$.

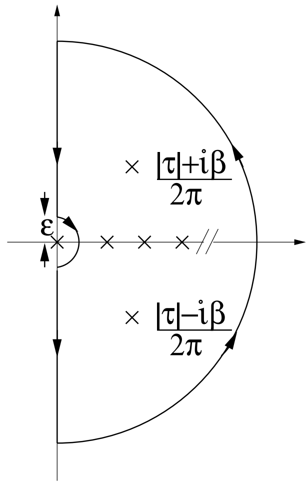


Figure 11.4.4: Contour used in Example 11.4.7.

From the residue theorem,

$$\oint_C \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz$$

$$= 2\pi i \sum_{n=1}^{\infty} \operatorname{Res}\left\{\cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right];n\right\}$$

$$+ 2\pi i \operatorname{Res}\left\{\cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right];\frac{|\tau|+\beta i}{2\pi}\right\}$$

$$+ 2\pi i \operatorname{Res}\left\{\cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right];\frac{|\tau|-\beta i}{2\pi}\right\}. \quad (\mathbf{11.4.41})$$

Now

$$\operatorname{Res}\left\{\cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right];n\right\}$$

$$= \lim_{z\to n}\frac{(z-n)\cos(\pi z)}{\sin(\pi z)}\lim_{z\to n}\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right] \quad (\mathbf{11.4.42})$$

$$= \frac{1}{\pi}\left[\frac{e^{-nc}}{(\tau+2n\pi)^2+\beta^2}+\frac{e^{-nc}}{(\tau-2n\pi)^2+\beta^2}\right], \quad (\mathbf{11.4.43})$$

$$\operatorname{Res}\left\{\cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right];\frac{|\tau|+\beta i}{2\pi}\right\}$$

$$= \lim_{z\to(|\tau|+\beta i)/2\pi}\frac{\cot(\pi z)}{4\pi^2}\left[\frac{(z-|\tau|-\beta i)e^{-cz}}{(z+\tau/2\pi)^2+\beta^2/4\pi^2}+\frac{(z-|\tau|-\beta i)e^{-cz}}{(z-\tau/2\pi)^2+\beta^2/4\pi^2}\right] \quad (\mathbf{11.4.44})$$

$$= \frac{\cot(|\tau|/2+\beta i/2)\exp(-c|\tau|/2\pi)[\cos(c\beta/2\pi)-i\sin(c\beta/2\pi)]}{4\pi\beta i}, \quad (\mathbf{11.4.45})$$

and

$$\operatorname{Res}\left\{\cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right];\frac{|\tau|-\beta i}{2\pi}\right\}$$

$$= \lim_{z\to(|\tau|-\beta i)/2\pi} \frac{\cot(\pi z)}{4\pi^2}\left[\frac{(z-|\tau|+\beta i)e^{-cz}}{(z+\tau/2\pi)^2+\beta^2/4\pi^2}+\frac{(z-|\tau|+\beta i)e^{-cz}}{(z-\tau/2\pi)^2+\beta^2/4\pi^2}\right] \quad (11.4.46)$$

$$= \frac{\cot(|\tau|/2-\beta i/2)\exp(-c|\tau|/2\pi)[\cos(c\beta/2\pi)+i\sin(c\beta/2\pi)]}{-4\pi\beta i}. \quad (11.4.47)$$

Therefore,

$$\oint_C \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz$$

$$= 2i\sum_{n=1}^{\infty}\left[\frac{e^{-nc}}{(\tau+2n\pi)^2+\beta^2}+\frac{e^{-nc}}{(\tau-2n\pi)^2+\beta^2}\right]$$
$$+\frac{i}{2\beta}\frac{e^{i|\tau|}+e^{\beta}}{e^{i|\tau|}-e^{\beta}}e^{-c|\tau|/2\pi}[\cos(c\beta/2\pi)-i\sin(c\beta/2\pi)]$$
$$-\frac{i}{2\beta}\frac{e^{i|\tau|}+e^{-\beta}}{e^{i|\tau|}-e^{-\beta}}e^{-c|\tau|/2\pi}[\cos(c\beta/2\pi)+i\sin(c\beta/2\pi)] \quad (11.4.48)$$

$$= 2i\sum_{n=1}^{\infty}\left[\frac{e^{-nc}}{(\tau+2n\pi)^2+\beta^2}+\frac{e^{-nc}}{(\tau-2n\pi)^2+\beta^2}\right]$$
$$-\frac{i}{\beta}\frac{\sinh(\beta)\cos(c\beta/2\pi)+\sin(|\tau|)\sin(c\beta/2\pi)}{\cosh(\beta)-\cos(\tau)}e^{-c|\tau|/2\pi}, \quad (11.4.49)$$

where $\cot(\alpha) = i(e^{2i\alpha}+1)/(e^{2i\alpha}-1)$, and we made extensive use of Euler's formula.

Let us now evaluate the contour integral by direct integration. The contribution from the integration along the semicircle at infinity vanishes according to Jordan's lemma. Indeed, that is why this particular contour was chosen. Therefore,

$$\oint_C \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz$$

$$= \int_{i\infty}^{i\epsilon} \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz$$
$$+\int_{C_\epsilon} \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz$$
$$+\int_{-i\epsilon}^{-i\infty} \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz. \quad (11.4.50)$$

Now, because $z = iy$,

$$\int_{i\infty}^{i\epsilon} \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz$$

$$= \int_{\infty}^{\epsilon} \coth(\pi y)\left[\frac{e^{-icy}}{(\tau+2\pi iy)^2+\beta^2}+\frac{e^{-icy}}{(\tau-2\pi iy)^2+\beta^2}\right]dy \quad (11.4.51)$$

$$= -2\int_{\epsilon}^{\infty} \frac{\coth(\pi y)(\tau^2+\beta^2-4\pi^2y^2)e^{-icy}}{(\tau^2+\beta^2-4\pi^2y^2)^2+16\pi^2\tau^2y^2}\,dy, \quad (11.4.52)$$

$$\int_{-i\epsilon}^{-i\infty} \cot(\pi z)\left[\frac{e^{-cz}}{(\tau+2\pi z)^2+\beta^2}+\frac{e^{-cz}}{(\tau-2\pi z)^2+\beta^2}\right]dz$$

$$= \int_{-\epsilon}^{-\infty} \coth(\pi y) \left[\frac{e^{-icy}}{(\tau + 2\pi i y)^2 + \beta^2} + \frac{e^{-icy}}{(\tau - 2\pi i y)^2 + \beta^2} \right] dy \quad (\mathbf{11.4.53})$$

$$= 2 \int_{\epsilon}^{\infty} \frac{\coth(\pi y)(\tau^2 + \beta^2 - 4\pi^2 y^2) e^{icy}}{(\tau^2 + \beta^2 - 4\pi^2 y^2)^2 + 16\pi^2 \tau^2 y^2} \, dy, \quad (\mathbf{11.4.54})$$

and

$$\int_{C_\epsilon} \cot(\pi z) \left[\frac{e^{-cz}}{(\tau + 2\pi z)^2 + \beta^2} + \frac{e^{-cz}}{(\tau - 2\pi z)^2 + \beta^2} \right] dz$$
$$= \int_{\pi/2}^{-\pi/2} \left[\frac{1}{\pi \epsilon e^{\theta i}} - \frac{\pi \epsilon e^{\theta i}}{3} - \cdots \right] \epsilon i e^{\theta i} \, d\theta$$
$$\times \left[\frac{\exp(-c\epsilon e^{\theta i})}{(\tau + 2\pi \epsilon e^{\theta i})^2 + \beta^2} + \frac{\exp(-c\epsilon e^{\theta i})}{(\tau - 2\pi \epsilon e^{\theta i})^2 + \beta^2} \right]. \quad (\mathbf{11.4.55})$$

In the limit of $\epsilon \to 0$,

$$\oint_C \cot(\pi z) \left[\frac{e^{-cz}}{(\tau + 2\pi z)^2 + \beta^2} + \frac{e^{-cz}}{(\tau - 2\pi z)^2 + \beta^2} \right] dz$$
$$= 4i \int_0^{\infty} \frac{\coth(\pi y)(\tau^2 + \beta^2 - 4\pi^2 y^2) \sin(cy)}{(\tau^2 + \beta^2 - 4\pi^2 y^2)^2 + 16\pi^2 \tau^2 y^2} \, dy - \frac{2i}{\tau^2 + \beta^2} \quad (\mathbf{11.4.56})$$
$$= 2i \sum_{n=1}^{\infty} \left[\frac{e^{-nc}}{(\tau + 2n\pi)^2 + \beta^2} + \frac{e^{-nc}}{(\tau - 2n\pi)^2 + \beta^2} \right]$$
$$- \frac{i}{\beta} \frac{\sinh(\beta)\cos(c\beta/2\pi) + \sin(|\tau|)\sin(c\beta/2\pi)}{\cosh(\beta) - \cos(\tau)} e^{-c|\tau|/2\pi}, \quad (\mathbf{11.4.57})$$

or

$$4 \int_0^{\infty} \frac{\coth(\pi y)(\tau^2 + \beta^2 - 4\pi^2 y^2) \sin(cy)}{(\tau^2 + \beta^2 - 4\pi^2 y^2)^2 + 16\pi^2 \tau^2 y^2} \, dy$$
$$= 2 \sum_{n=1}^{\infty} \left[\frac{e^{-nc}}{(\tau + 2n\pi)^2 + \beta^2} + \frac{e^{-nc}}{(\tau - 2n\pi)^2 + \beta^2} \right] \quad (\mathbf{11.4.58})$$
$$- \frac{1}{\beta} \frac{\sinh(\beta)\cos(c\beta/2\pi) + \sin(|\tau|)\sin(c\beta/2\pi)}{\cosh(\beta) - \cos(\tau)} e^{-c|\tau|/2\pi} + \frac{2}{\tau^2 + \beta^2}.$$

If we let $y = x/2\pi$,

$$\frac{\beta}{\pi} \int_0^{\infty} \frac{\coth(x/2)(\tau^2 + \beta^2 - x^2) \sin(cx/2\pi)}{(\tau^2 + \beta^2 - x^2)^2 + 4\tau^2 x^2} \, dx$$
$$= 2\beta \sum_{n=1}^{\infty} \left[\frac{e^{-nc}}{(\tau + 2n\pi)^2 + \beta^2} + \frac{e^{-nc}}{(\tau - 2n\pi)^2 + \beta^2} \right] \quad (\mathbf{11.4.59})$$
$$- \frac{\sinh(\beta)\cos(c\beta/2\pi) + \sin(|\tau|)\sin(c\beta/2\pi)}{\cosh(\beta) - \cos(\tau)} e^{-c|\tau|/2\pi} + \frac{2\beta}{\tau^2 + \beta^2}.$$

□

Problems

1. Use direct integration to find the inverse of the Fourier transform $F(\omega) = i\omega\pi e^{-|\omega|}/2$. Check your answer using MATLAB.

Use partial fractions to invert the following Fourier transforms:

2. $\dfrac{1}{(1+i\omega)(1+2i\omega)}$

3. $\dfrac{1}{(1+i\omega)(1-i\omega)}$

4. $\dfrac{i\omega}{(1+i\omega)(1+2i\omega)}$

5. $\dfrac{1}{(1+i\omega)(1+2i\omega)^2}$

Then check your answer using MATLAB.

By taking the appropriate closed contour, find the inverse of the following Fourier transforms by contour integration. The parameter a is real and positive.

6. $\dfrac{1}{\omega^2+a^2}$

7. $\dfrac{\omega}{\omega^2+a^2}$

8. $\dfrac{\omega}{(\omega^2+a^2)^2}$

9. $\dfrac{\omega^2}{(\omega^2+a^2)^2}$

10. $\dfrac{1}{\omega^2-3i\omega-3}$

11. $\dfrac{1}{(\omega-ia)^{2n+2}}$

12. $\dfrac{\omega^2}{(\omega^2-1)^2+4a^2\omega^2}$

13. $\dfrac{3}{(2-\omega i)(1+\omega i)}$

Then check your answer using MATLAB.

14. Find the inverse of $F(\omega) = \cos(\omega)/(\omega^2+a^2)$, $a>0$, by first rewriting the transform as

$$F(\omega) = \frac{e^{i\omega}}{2(\omega^2+a^2)} + \frac{e^{-i\omega}}{2(\omega^2+a^2)}$$

and then using the residue theorem on each term.

15. Find[12] the inverse Fourier transform for

$$F_\pm(\omega) = \frac{e^{\pm i\omega}}{(\omega-ai)\left(R^2e^{\omega i}-e^{-\omega i}\right)} = \frac{e^{\pm i\omega-i\omega}}{(\omega-ai)\left(R^2-e^{-2\omega i}\right)},$$

where $a>0$ and $R>1$. Hint: You must find separate inverses for different time intervals. For example, in the case of $F_+(\omega)$, you must examine the special cases of $t<0$ and $t>0$.

16. As we shall show shortly, Fourier transforms can be used to solve differential equations. During the solution of the heat equation, Taitel et al.[13] inverted the Fourier transform

$$F(\omega) = \frac{\cosh(y\sqrt{\omega^2+1}\,)}{\sqrt{\omega^2+1}\,\sinh(p\sqrt{\omega^2+1}/2)},$$

[12] See Scharstein, R. W., 1992: Transient electromagnetic plane wave reflection from a dielectric slab. *IEEE Trans. Educ.*, **35**, 170–175.

[13] Taitel, Y., M. Bentwich, and A. Tamir, 1973: Effects of upstream and downstream boundary conditions on heat (mass) transfer with axial diffusion. *Int. J. Heat Mass Transfer*, **16**, 359–369.

where y and p are real. Show that they should have found

$$f(t) = \frac{e^{-|t|}}{p} + \frac{2}{p}\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{1+4n^2\pi^2/p^2}} \cos\left(\frac{2n\pi y}{p}\right) e^{-\sqrt{1+4n^2\pi^2/p^2}\,|t|}.$$

In this case, our time variable t was their spatial variable $x - \xi$.

17. Find the inverse of the Fourier transform

$$F(\omega) = \left[\cos\left\{\frac{\omega L}{\beta[1 + i\gamma\ \mathrm{sgn}(\omega)]}\right\}\right]^{-1},$$

where L, β, and γ are real and positive and $\mathrm{sgn}(z) = 1$ if $\Re(z) > 0$ and -1 if $\Re(z) < 0$.

Use the residue theorem to verify the following integrals:

18. $\displaystyle\int_{-\infty}^{\infty} \frac{\sin(x)}{x^2 + 4x + 5}\, dx = -\frac{\pi}{e}\sin(2)$

19. $\displaystyle\int_0^{\infty} \frac{\cos(x)}{(x^2+1)^2}\, dx = \frac{\pi}{2e}$

20. $\displaystyle\int_{-\infty}^{\infty} \frac{x\sin(ax)}{x^2 + 4}\, dx = \pi e^{-2a}$

21. $\displaystyle\int_0^{\infty} \frac{x^2\cos(ax)}{(x^2+b^2)^2}\, dx = \frac{\pi}{4b}(1 - ab)e^{-ab}$

where $a, b > 0$.

22. The concept of forced convection is normally associated with heat streaming through a duct or past an obstacle. Bentwich[14] showed that a similar transport can exist when convection results from a wave traveling through an essentially stagnant fluid. In the process of computing the amount of heating, he proved the following identity:

$$\int_{-\infty}^{\infty} \frac{\cosh(hx) - 1}{x\,\sinh(hx)} \cos(ax)\, dx = \ln[\coth(|a|\pi/h)], \qquad h > 0.$$

Confirm his result.

11.5 CONVOLUTION

The most important property of Fourier transforms is convolution. We shall use it extensively in the solution of differential equations and the design of filters because it yields in time or space the effect of multiplying two transforms together.

The convolution operation is

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(x)g(t-x)\, dx = \int_{-\infty}^{\infty} f(t-x)g(x)\, dx. \qquad \mathbf{(11.5.1)}$$

Then,

$$\mathcal{F}[f(t) * g(t)] = \int_{-\infty}^{\infty} f(x)e^{-i\omega x}\left[\int_{-\infty}^{\infty} g(t-x)e^{-i\omega(t-x)}dt\right]dx \qquad \mathbf{(11.5.2)}$$

$$= \int_{-\infty}^{\infty} f(x)G(\omega)e^{-i\omega x}dx = F(\omega)G(\omega). \qquad \mathbf{(11.5.3)}$$

---

[14] Bentwich, M., 1966: Convection enforced by surface and tidal waves. *Int. J. Heat Mass Transfer*, **9**, 663–670.

Thus, the Fourier transform of the convolution of two functions equals the product of the Fourier transforms of each of the functions.

• Example 11.5.1

Let us verify the convolution theorem using the functions $f(t) = H(t+a) - H(t-a)$ and $g(t) = e^{-t}H(t)$, where $a > 0$.

The convolution of $f(t)$ with $g(t)$ is

$$f(t) * g(t) = \int_{-\infty}^{\infty} e^{-(t-x)} H(t-x) \left[H(x+a) - H(x-a)\right] dx = e^{-t} \int_{-a}^{a} e^{x} H(t-x)\, dx. \tag{11.5.4}$$

If $t < -a$, then the integrand of Equation 11.5.4 is always zero and $f(t) * g(t) = 0$. If $t > a$,

$$f(t) * g(t) = e^{-t} \int_{-a}^{a} e^{x}\, dx = e^{-(t-a)} - e^{-(t+a)}. \tag{11.5.5}$$

Finally, for $-a < t < a$,

$$f(t) * g(t) = e^{-t} \int_{-a}^{t} e^{x}\, dx = 1 - e^{-(t+a)}. \tag{11.5.6}$$

In summary,

$$f(t) * g(t) = \begin{cases} 0, & t \le -a, \\ 1 - e^{-(t+a)}, & -a \le t \le a, \\ e^{-(t-a)} - e^{-(t+a)}, & a \le t. \end{cases} \tag{11.5.7}$$

□

As an alternative to examining various cases involving the value of t, we could have used MATLAB to evaluate Equation 11.5.4. The MATLAB instructions are as follows:

```
>> syms f t x
>> syms a positive
>> f = 'exp(x-t)*Heaviside(t-x)*(Heaviside(x+a)-Heaviside(x-a))'
>> int(f,x,-inf,inf)
```

This yields

```
ans =
Heaviside(t+a)-Heaviside(t+a)*exp(-a-t)
    -Heaviside(t-a)+Heaviside(t-a)*exp(a-t)
```

The Fourier transform of $f(t) * g(t)$ is

$$\mathcal{F}[f(t) * g(t)] = \int_{-a}^{a} \left[1 - e^{-(t+a)}\right] e^{-i\omega t} dt + \int_{a}^{\infty} \left[e^{-(t-a)} - e^{-(t+a)}\right] e^{-i\omega t} dt \tag{11.5.8}$$

$$= \frac{2\sin(\omega a)}{\omega} - \frac{2i\sin(\omega a)}{1+\omega i} = \frac{2\sin(\omega a)}{\omega}\left(\frac{1}{1+\omega i}\right) = F(\omega)G(\omega) \tag{11.5.9}$$

and the convolution theorem is true for this special case. The Fourier transform Equation 11.5.9 could also be obtained by substituting our earlier MATLAB result into `fourier` and then using `simplify(ans)`.

• **Example 11.5.2**

Let us consider the convolution of $f(t) = f_+(t)H(t)$ with $g(t) = g_+H(t)$. Note that both of the functions are nonzero only for $t > 0$.

From the definition of convolution,

$$f(t) * g(t) = \int_{-\infty}^{\infty} f_+(t-x)H(t-x)g_+(x)H(x)\,dx = \int_0^{\infty} f_+(t-x)H(t-x)g_+(x)\,dx. \tag{11.5.10}$$

For $t < 0$, the integrand is always zero and $f(t) * g(t) = 0$. For $t > 0$,

$$f(t) * g(t) = \int_0^t f_+(t-x)g_+(x)\,dx. \tag{11.5.11}$$

Therefore, in general,

$$f(t) * g(t) = \left[\int_0^t f_+(t-x)g_+(x)\,dx\right] H(t). \tag{11.5.12}$$

This is the definition of convolution that we will use for Laplace transforms where all of the functions equal zero for $t < 0$. □

The convolution operation also applies to Fourier transforms, in what is commonly known as *frequency convolution*. We now prove that

$$\mathcal{F}[f(t)g(t)] = \frac{F(\omega) * G(\omega)}{2\pi}, \tag{11.5.13}$$

where

$$F(\omega) * G(\omega) = \int_{-\infty}^{\infty} F(\tau)G(\omega-\tau)\,d\tau, \tag{11.5.14}$$

where $F(\omega)$ and $G(\omega)$ are the Fourier transforms of $f(t)$ and $g(t)$, respectively.

Proof: Starting with

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\tau)e^{i\tau t}\,d\tau, \tag{11.5.15}$$

we can multiply the inverse of $F(\tau)$ by $g(t)$ so that we obtain

$$f(t)g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\tau)g(t)e^{i\tau t}\,d\tau. \tag{11.5.16}$$

Then, taking the Fourier transform of Equation 11.5.16, we find that

$$\mathcal{F}[f(t)g(t)] = \int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} F(\tau)g(t)e^{i\tau t}\,d\tau\right] e^{-i\omega t}\,dt \tag{11.5.17}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\tau)\left[\int_{-\infty}^{\infty} g(t)e^{-i(\omega-\tau)t}\,dt\right] d\tau \tag{11.5.18}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\tau)G(\omega-\tau)\,d\tau = \frac{F(\omega) * G(\omega)}{2\pi}. \tag{11.5.19}$$

Thus, the multiplication of two functions in the time domain is equivalent to the convolution of their spectral densities in the frequency domain. □

Problems

1. Show that $e^{-t}H(t) * e^{-t}H(t) = te^{-t}H(t)$. Then verify your result using MATLAB.

2. Show that $e^{-t}H(t) * e^{t}H(-t) = \frac{1}{2}e^{-|t|}$. Then verify your result using MATLAB.

3. Show that $e^{-t}H(t) * e^{-2t}H(t) = \left(e^{-t} - e^{-2t}\right)H(t)$. Then verify your result using MATLAB.

4. Show that

$$e^{t}H(-t) * [H(t) - H(t-2)] = \begin{cases} e^{t} - e^{t-2}, & t \le 0, \\ 1 - e^{t-2}, & 0 \le t \le 2, \\ 0, & 2 \le t. \end{cases}$$

Then verify your result using MATLAB.

5. Show that

$$[H(t) - H(t-2)] * [H(t) - H(t-2)] = \begin{cases} 0, & t \le 0, \\ t, & 0 \le t \le 2, \\ 4 - t, & 2 \le t \le 4, \\ 0, & 4 \le t. \end{cases}$$

Then try and verify your result using MATLAB. What do you have to do to make it work?

6. Show that $e^{-|t|} * e^{-|t|} = (1 + |t|)e^{-|t|}$.

7. Prove that the convolution of two Dirac delta functions is a Dirac delta function.

11.6 THE SOLUTION OF ORDINARY DIFFERENTIAL EQUATIONS BY FOURIER TRANSFORMS

As with Laplace transforms, we may use Fourier transforms to solve ordinary differential equations. However, this method gives only the particular solution and we must find the complementary solution separately.

Consider the differential equation

$$y' + y = \tfrac{1}{2}e^{-|t|}, \qquad -\infty < t < \infty. \tag{11.6.1}$$

Taking the Fourier transform of both sides of Equation 11.6.1,

$$i\omega Y(\omega) + Y(\omega) = \frac{1}{\omega^2 + 1}, \tag{11.6.2}$$

where we used the derivative rule, Equation 11.3.17, to obtain the transform of y' and $Y(\omega) = \mathcal{F}[y(t)]$. Therefore,

$$Y(\omega) = \frac{1}{(\omega^2 + 1)(1 + \omega i)}. \tag{11.6.3}$$

Applying the inversion integral to Equation 11.6.3,

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{it\omega}}{(\omega^2+1)(1+\omega i)}\, d\omega. \tag{11.6.4}$$

We evaluate Equation 11.6.4 by contour integration. For $t > 0$ we close the line integral with an infinite semicircle in the upper half of the ω-plane. The integration along this arc equals zero by Jordan's lemma. Within this closed contour we have a second-order pole at $z = i$. Therefore,

$$\operatorname{Res}\left[\frac{e^{itz}}{(z^2+1)(1+zi)}; i\right] = \lim_{z\to i} \frac{d}{dz}\left[(z-i)^2 \frac{e^{itz}}{i(z-i)^2(z+i)}\right] = \frac{te^{-t}}{2i} + \frac{e^{-t}}{4i} \tag{11.6.5}$$

and

$$y(t) = \frac{1}{2\pi}(2\pi i)\left[\frac{te^{-t}}{2i} + \frac{e^{-t}}{4i}\right] = \frac{e^{-t}}{4}(2t+1). \tag{11.6.6}$$

For $t < 0$, we again close the line integral with an infinite semicircle but this time it is in the lower half of the ω-plane. The contribution from the line integral along the arc vanishes by Jordan's lemma. Within the contour, we have a simple pole at $z = -i$. Therefore,

$$\operatorname{Res}\left[\frac{e^{itz}}{(z^2+1)(1+zi)}; -i\right] = \lim_{z\to -i} (z+i) \frac{e^{itz}}{i(z+i)(z-i)^2} = -\frac{e^{t}}{4i}, \tag{11.6.7}$$

and

$$y(t) = \frac{1}{2\pi}(-2\pi i)\left(-\frac{e^{t}}{4i}\right) = \frac{e^{t}}{4}. \tag{11.6.8}$$

The minus sign in front of the $2\pi i$ results from the contour being taken in the clockwise direction or negative sense. Using the step function, we can combine Equation 11.6.6 and Equation 11.6.8 into the single expression

$$y(t) = \tfrac{1}{4}e^{-|t|} + \tfrac{1}{2}te^{-t}H(t). \tag{11.6.9}$$

Note that we only found the particular or forced solution to Equation 11.6.1. The most general solution therefore requires that we add the complementary solution Ae^{-t}, yielding

$$y(t) = Ae^{-t} + \tfrac{1}{4}e^{-|t|} + \tfrac{1}{2}te^{-t}H(t). \tag{11.6.10}$$

The arbitrary constant A would be determined by the initial condition, which we have not specified.

We could also have solved this problem using MATLAB. The MATLAB script

```
clear
% define symbolic variables
syms omega t Y
% take Fourier transform of left side of differential equation
LHS = fourier(diff(sym('y(t)'))+sym('y(t)'),t,omega);
% take Fourier transform of right side of differential equation
RHS = fourier(1/2*exp(-abs(t)),t,omega);
% set Y for Fourier transform of y
%     and introduce initial conditions
newLHS = subs(LHS,'fourier(y(t),t,omega)',Y);
```

```
% solve for Y
Y = solve(newLHS-RHS,Y);
% invert Fourier transform and find y(t)
y = ifourier(Y,omega,t)
```

yields

```
y =
1/4*exp(t)*Heaviside(-t)+1/2*exp(-t)*t*Heaviside(t)
    +1/4*exp(-t)*Heaviside(t)
```

which is equivalent to Equation 11.6.9.

Consider now a more general problem of

$$y' + y = f(t), \qquad -\infty < t < \infty, \tag{11.6.11}$$

where we assume that $f(t)$ has the Fourier transform $F(\omega)$. Then the Fourier-transformed solution to Equation 11.6.11 is

$$Y(\omega) = \frac{1}{1+\omega i} F(\omega) = G(\omega)F(\omega) \qquad \text{or} \qquad y(t) = g(t) * f(t), \tag{11.6.12}$$

where $g(t) = \mathcal{F}^{-1}[1/(1+\omega i)] = e^{-t}H(t)$. Thus, we can obtain our solution in one of two ways. First, we can take the Fourier transform of $f(t)$, multiply this transform by $G(\omega)$, and finally compute the inverse. The second method requires a convolution of $f(t)$ with $g(t)$. Which method is easiest depends upon $f(t)$ and $g(t)$.

In summary, we can use Fourier transforms to find particular solutions to differential equations. The complete solution consists of this particular solution plus any homogeneous solution that we need to satisfy the initial conditions. Convolution of the Green's function with the forcing function also gives the particular solution.

Problems

Find the particular solutions for the following differential equations. For Problems 1–3, verify your solution using MATLAB.

1. $y'' + 3y' + 2y = e^{-t}H(t)$
2. $y'' + 4y' + 4y = \frac{1}{2}e^{-|t|}$
3. $y'' - 4y' + 4y = e^{-t}H(t)$
4. $y^{iv} - \lambda^4 y = \delta(x)$,

where λ has a positive real part and a negative imaginary part.

11.7 THE SOLUTION OF LAPLACE'S EQUATION ON THE UPPER HALF-PLANE

In this section we shall use Fourier integrals and convolution to find the solution of Laplace's equation on the upper half-plane $y > 0$. We require that the solution remains bounded over the entire domain and specify it along the x-axis, $u(x, 0) = f(x)$. Under these conditions, we can take the Fourier transform of Laplace's equation and find that

$$\int_{-\infty}^{\infty} \frac{\partial^2 u}{\partial x^2} e^{-i\omega x}\, dx + \int_{-\infty}^{\infty} \frac{\partial^2 u}{\partial y^2} e^{-i\omega x}\, dx = 0. \tag{11.7.1}$$

If everything is sufficiently differentiable, we may successively integrate by parts the first integral in Equation 11.7.1, which yields

$$\int_{-\infty}^{\infty} \frac{\partial^2 u}{\partial x^2} e^{-i\omega x}\,dx = \frac{\partial u}{\partial x} e^{-i\omega x}\Big|_{-\infty}^{\infty} + i\omega \int_{-\infty}^{\infty} \frac{\partial u}{\partial x} e^{-i\omega x}\,dx \tag{11.7.2}$$

$$= i\omega\, u(x,y) e^{-i\omega x}\big|_{-\infty}^{\infty} - \omega^2 \int_{-\infty}^{\infty} u(x,y) e^{-i\omega x}\,dx \tag{11.7.3}$$

$$= -\omega^2 U(\omega, y), \tag{11.7.4}$$

where

$$U(\omega, y) = \int_{-\infty}^{\infty} u(x,y) e^{-i\omega x}\,dx. \tag{11.7.5}$$

The second integral becomes

$$\int_{-\infty}^{\infty} \frac{\partial^2 u}{\partial y^2} e^{-i\omega x}\,dx = \frac{d^2}{dy^2}\left[\int_{-\infty}^{\infty} u(x,y) e^{-i\omega x}\,dx\right] = \frac{d^2 U(\omega,y)}{dy^2}, \tag{11.7.6}$$

along with the boundary condition that

$$F(\omega) = U(\omega, 0) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x}\,dx. \tag{11.7.7}$$

Consequently, we reduced Laplace's equation, a partial differential equation, to an ordinary differential equation in y, where ω is merely a parameter:

$$\frac{d^2 U(\omega,y)}{dy^2} - \omega^2 U(\omega, y) = 0, \tag{11.7.8}$$

with the boundary condition $U(\omega, 0) = F(\omega)$. The solution to Equation 11.7.8 is

$$U(\omega, y) = A(\omega) e^{|\omega| y} + B(\omega) e^{-|\omega| y}, \quad 0 \le y. \tag{11.7.9}$$

We must discard the $e^{|\omega|y}$ term because it becomes unbounded as we go to infinity along the y-axis. The boundary condition results in $B(\omega) = F(\omega)$. Consequently,

$$U(\omega, y) = F(\omega) e^{-|\omega| y}. \tag{11.7.10}$$

The inverse of the Fourier transform $e^{-|\omega|y}$ equals

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-|\omega|y} e^{i\omega x}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{0} e^{\omega y} e^{i\omega x}\,d\omega + \frac{1}{2\pi}\int_{0}^{\infty} e^{-\omega y} e^{i\omega x}\,d\omega \tag{11.7.11}$$

$$= \frac{1}{2\pi}\int_{0}^{\infty} e^{-\omega y} e^{-i\omega x}\,d\omega + \frac{1}{2\pi}\int_{0}^{\infty} e^{-\omega y} e^{i\omega x}\,d\omega \tag{11.7.12}$$

$$= \frac{1}{\pi}\int_{0}^{\infty} e^{-\omega y}\cos(\omega x)\,d\omega \tag{11.7.13}$$

$$= \frac{1}{\pi}\left\{\frac{\exp(-\omega y)}{x^2+y^2}\left[-y\cos(\omega x) + x\sin(\omega x)\right]\right\}\Bigg|_{0}^{\infty} \tag{11.7.14}$$

$$= \frac{1}{\pi}\frac{y}{x^2+y^2}. \tag{11.7.15}$$

Furthermore, because Equation 11.7.10 is a convolution of two Fourier transforms, its inverse is

$$u(x,y) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{yf(t)}{(x-t)^2+y^2}\,dt. \tag{11.7.16}$$

Equation 11.7.16 is *Poisson's integral formula*[15] for the half-plane $y > 0$ or *Schwarz' integral formula*.[16]

• **Example 11.7.1**

As an example, let $u(x,0) = 1$ if $|x| < 1$ and $u(x,0) = 0$ otherwise. Then,

$$u(x,y) = \frac{1}{\pi}\int_{-1}^{1} \frac{y}{(x-t)^2+y^2}\,dt = \frac{1}{\pi}\left[\tan^{-1}\left(\frac{1-x}{y}\right) + \tan^{-1}\left(\frac{1+x}{y}\right)\right]. \tag{11.7.17}$$

Problems

Find the solution to Laplace's equation in the upper half-plane for the following boundary conditions:

1. $u(x,0) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$

2. $u(x,0) = \begin{cases} 1, & x > 0, \\ -1, & x < 0. \end{cases}$

3. $u(x,0) = \begin{cases} T_0, & x < 0, \\ 0, & x > 0. \end{cases}$

4. $u(x,0) = \begin{cases} 2T_0, & x < -1, \\ T_0, & -1 < x < 1, \\ 0, & 1 < x. \end{cases}$

5. $u(x,0) = \begin{cases} T_0, & -1 < x < 0, \\ T_0 + (T_1 - T_0)x, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$

6. $u(x,0) = \begin{cases} T_0, & x < a_1, \\ T_1, & a_1 < x < a_2, \\ T_2, & a_2 < x < a_3, \\ \vdots & \vdots \\ T_n, & a_n < x. \end{cases}$

11.8 THE SOLUTION OF THE HEAT EQUATION

We now consider the problem of one-dimensional heat flow in a rod of infinite length with insulated sides. Although there are no boundary conditions because the slab is of infinite extent, we do require that the solution remains bounded as we go to either positive or negative infinity. The initial temperature within the rod is $u(x,0) = f(x)$.

Employing the product solution technique of Section 8.3, we begin by assuming that $u(x,t) = X(x)T(t)$ with

$$T' + a^2\lambda T = 0, \tag{11.8.1}$$

[15] Poisson, S. D., 1823: Suite du mémoire sur les intégrales définies et sur la sommation des séries. *J. École Polytech.*, **19**, 404–509. See pg. 462.

[16] Schwarz, H. A., 1870: Über die Integration der partiellen Differentialgleichung $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = 0$ für die Fläche eines Kreises. *Vierteljahrsschr. Naturforsch. Ges. Zürich*, **15**, 113–128.

and

$$X'' + \lambda X = 0. \tag{11.8.2}$$

Solutions to Equation 11.8.1 and Equation 11.8.2, which remain finite over the entire x-domain, are

$$X(x) = E\cos(kx) + F\sin(kx), \tag{11.8.3}$$

and

$$T(t) = C\exp(-k^2a^2t). \tag{11.8.4}$$

Because we do not have any boundary conditions, we must include *all* possible values of k. Thus, when we sum all of the product solutions according to the principle of linear superposition, we obtain the integral

$$u(x,t) = \int_0^\infty [A(k)\cos(kx) + B(k)\sin(kx)]e^{-k^2a^2t}\,dk. \tag{11.8.5}$$

We can satisfy the initial condition by choosing

$$A(k) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x)\cos(kx)\,dx, \tag{11.8.6}$$

and

$$B(k) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x)\sin(kx)\,dx, \tag{11.8.7}$$

because the initial condition has the form of a Fourier integral

$$f(x) = \int_0^\infty [A(k)\cos(kx) + B(k)\sin(kx)]\,dk, \tag{11.8.8}$$

when $t = 0$.

Several important results follow by rewriting Equation 11.8.8 as

$$u(x,t) = \frac{1}{\pi}\int_0^\infty \left[\int_{-\infty}^{\infty} f(\xi)\cos(k\xi)\cos(kx)\,d\xi + \int_{-\infty}^{\infty} f(\xi)\sin(k\xi)\sin(kx)\,d\xi\right]e^{-k^2a^2t}\,dk. \tag{11.8.9}$$

Combining terms,

$$u(x,t) = \frac{1}{\pi}\int_0^\infty \left\{\int_{-\infty}^{\infty} f(\xi)[\cos(k\xi)\cos(kx) + \sin(k\xi)\sin(kx)]\,d\xi\right\}e^{-k^2a^2t}\,dk \tag{11.8.10}$$

$$= \frac{1}{\pi}\int_0^\infty \left[\int_{-\infty}^{\infty} f(\xi)\cos[k(\xi - x)]\,d\xi\right]e^{-k^2a^2t}\,dk. \tag{11.8.11}$$

Reversing the order of integration,

$$u(x,t) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(\xi)\left[\int_0^\infty \cos[k(\xi - x)]e^{-k^2a^2t}\,dk\right]d\xi. \tag{11.8.12}$$

The inner integral is called the *source function*. We may compute its value through an integration on the complex plane; it equals

$$\int_0^\infty \cos[k(\xi - x)]\exp(-k^2a^2t)\,dk = \left(\frac{\pi}{4a^2t}\right)^{1/2}\exp\left[-\frac{(\xi - x)^2}{4a^2t}\right], \tag{11.8.13}$$

if $0 < t$. This gives the final form for the temperature distribution:

$$u(x,t) = \frac{1}{\sqrt{4a^2\pi t}} \int_{-\infty}^{\infty} f(\xi) \exp\left[-\frac{(\xi - x)^2}{4a^2 t}\right] d\xi. \tag{11.8.14}$$

• Example 11.8.1

Let us find the temperature field if the initial distribution is

$$u(x,0) = \begin{cases} T_0, & x > 0, \\ -T_0, & x < 0. \end{cases} \tag{11.8.15}$$

Then

$$u(x,t) = \frac{T_0}{\sqrt{4a^2\pi t}} \left\{ \int_0^{\infty} \exp\left[-\frac{(\xi - x)^2}{4a^2 t}\right] d\xi - \int_{-\infty}^{0} \exp\left[-\frac{(\xi - x)^2}{4a^2 t}\right] d\xi \right\} \tag{11.8.16}$$

$$= \frac{T_0}{\sqrt{\pi}} \left[\int_{-x/2a\sqrt{t}}^{\infty} e^{-\tau^2}\, d\tau - \int_{x/2a\sqrt{t}}^{\infty} e^{-\tau^2}\, d\tau \right] = \frac{T_0}{\sqrt{\pi}} \int_{-x/2a\sqrt{t}}^{x/2a\sqrt{t}} e^{-\tau^2}\, d\tau \tag{11.8.17}$$

$$= \frac{2T_0}{\sqrt{\pi}} \int_0^{x/2a\sqrt{t}} e^{-\tau^2}\, d\tau = T_0 \operatorname{erf}\left(\frac{x}{2a\sqrt{t}}\right), \tag{11.8.18}$$

where erf(·) is the error function. □

• Example 11.8.2: Kelvin's estimate of the age of the earth

In the middle of the nineteenth century, Lord Kelvin[17] estimated the age of the earth using the observed vertical temperature gradient at the earth's surface. He hypothesized that the earth was initially formed at a uniform high temperature T_0 and that its surface was subsequently maintained at the lower temperature of T_S. Assuming that most of the heat conduction occurred near the earth's surface, he reasoned that he could neglect the curvature of the earth, consider the earth's surface planar, and employ our one-dimensional heat conduction model in the vertical direction to compute the observed heat flux.

Following Kelvin, we model the earth's surface as a flat plane with an infinitely deep earth below ($z > 0$). Initially the earth has the temperature T_0. Suddenly we drop the temperature at the surface to T_S. We wish to find the heat flux across the boundary at $z = 0$ from the earth into an infinitely deep atmosphere.

The first step is to redefine our temperature scale $v(z,t) = u(z,t) + T_S$, where $v(z,t)$ is the observed temperature so that $u(0,t) = 0$ at the surface. Next, in order to use Equation 11.8.14, we must define our initial state for $z < 0$. To maintain the temperature $u(0,t) = 0$, the initial temperature field $f(z)$ must be an odd function, or

$$f(z) = \begin{cases} T_0 - T_S, & z > 0, \\ T_S - T_0, & z < 0. \end{cases} \tag{11.8.19}$$

[17] Thomson, W., 1863: On the secular cooling of the earth. *Philos. Mag., Ser. 4*, **25**, 157–170.

From Equation 11.8.14,

$$u(z,t) = \frac{T_0 - T_S}{\sqrt{4a^2\pi t}} \left\{ \int_0^\infty \exp\left[-\frac{(\xi - z)^2}{4a^2 t}\right] d\xi - \int_{-\infty}^0 \exp\left[-\frac{(\xi - z)^2}{4a^2 t}\right] d\xi \right\} \quad \textbf{(11.8.20)}$$

$$= (T_0 - T_S) \operatorname{erf}\left(\frac{z}{2a\sqrt{t}}\right), \quad \textbf{(11.8.21)}$$

following the work in the previous example.

The heat flux q at the surface $z = 0$ is obtained by differentiating Equation 11.8.21 according to Fourier's law and evaluating the result at $z = 0$:

$$q = -\kappa \frac{\partial v}{\partial z}\bigg|_{z=0} = \frac{\kappa(T_S - T_0)}{a\sqrt{\pi t}}. \quad \textbf{(11.8.22)}$$

The surface heat flux is infinite at $t = 0$ because of the sudden application of the temperature T_S at $t = 0$. After that time, the heat flux decreases with time. Consequently, the time t at which we have the temperature gradient $\partial v(0,t)/\partial z$ is

$$t = \frac{(T_0 - T_S)^2}{\pi a^2 [\partial v(0,t)/\partial z]^2}. \quad \textbf{(11.8.23)}$$

For the present near-surface thermal gradient of 25 K/km, $T_0 - T_S = 2000$ K, and $a^2 = 1$ mm^2/s, the age of the earth from Equation 11.8.23 is 65 million years.

Although Kelvin realized that this was a very rough estimate, his calculation showed that the earth had a finite age. This was in direct contradiction to the contemporary geological principle of *uniformitarianism*: that the earth's surface and upper crust had remained unchanged in temperature and other physical quantities for millions and millions of years. The resulting debate would rage throughout the latter half of the nineteenth century and feature such luminaries as Kelvin, Charles Darwin, Thomas Huxley, and Oliver Heaviside.[18] Eventually Kelvin's arguments would prevail and uniformitarianism would fade into history.

Today, Kelvin's estimate is of academic interest because of the discovery of radioactivity at the turn of the twentieth century. During the first half of the twentieth century, geologists assumed that the radioactivity was uniformly distributed around the globe and restricted to the upper few tens of kilometers of the crust. Using this model they would then use observed heat fluxes to compute the distribution of radioactivity within the solid earth.[19] Now we know that the interior of the earth is quite dynamic; the oceans and continents are mobile and interconnected according to the theory of plate tectonics. However, geophysicists still use measured surface heat fluxes to infer the interior[20] of the earth. □

• Example 11.8.3

So far we have shown how a simple application of separation of variables and the Fourier transform yields solutions to the heat equation over the semi-infinite interval $(0, \infty)$

---

[18] See Burchfield, J. D., 1975: *Lord Kelvin and the Age of the Earth*. Science History Publ., 260 pp.

[19] See Slichter, L. B., 1941: Cooling of the earth. *Bull. Geol. Soc. Am.*, **52**, 561–600.

[20] Sclater, J. G., C. Jaupart, and D. Galson, 1980: The heat flow through oceanic and continental crust and the heat loss of the earth. *Rev. Geophys. Space Phys.*, **18**, 269–311.

via Equation 11.8.5. Can we still use this technique for more complicated versions of the heat equation? The answer is yes but the procedure is more complicated. We illustrate it by solving[21]

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^3 u}{\partial t \partial x^2} + a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t, \tag{11.8.24}$$

subject to the boundary conditions

$$u(0,t) = f(t), \quad \lim_{x\to 0} |u(x,t)| < \infty, \qquad 0 < t, \tag{11.8.25}$$

$$\lim_{x\to\infty} u(x,t) \to 0, \lim_{x\to\infty} u_x(x,t) \to 0, \qquad 0 < t, \tag{11.8.26}$$

and the initial condition

$$u(x,0) = 0, \qquad 0 < x < \infty. \tag{11.8.27}$$

We begin by multiplying Equation 11.8.24 by $\sin(kx)$ and integrating over x from 0 to ∞:

$$\alpha \int_0^\infty u_{txx} \sin(kx)\, dx + a^2 \int_0^\infty u_{xx} \sin(kx)\, dx = \int_0^\infty u_t \sin(kx)\, dx. \tag{11.8.28}$$

Next, we integrate by parts. For example,

$$\int_0^\infty u_{xx} \sin(kx)\, dx = u_x \sin(kx)\Big|_0^\infty - k \int_0^\infty u_x \cos(kx)\, dx \tag{11.8.29}$$

$$= -k \int_0^\infty u_x \cos(kx)\, dx \tag{11.8.30}$$

$$= -ku(x,t)\cos(kx)\Big|_0^\infty - k^2 \int_0^\infty u(x,t)\sin(kx)\, dx \tag{11.8.31}$$

$$= kf(t) - k^2 U(k,t), \tag{11.8.32}$$

where

$$U(k,t) = \int_0^\infty u(x,t)\sin(kx)\, dx, \tag{11.8.33}$$

and the boundary conditions have been used to simplify Equation 11.8.29 and Equation 11.8.31. Equation 11.8.33 is the definition of the *Fourier sine transform*. It and its mathematical cousin, the *Fourier cosine transform* $\int_0^\infty u(x,t)\cos(kx)\, dx$, are analogous to the half-range sine and cosine expansions that appear in solving the heat equation over the finite interval $(0, L)$. The difference here is that our range runs from 0 to ∞.

Applying the same technique to the other terms, we obtain

$$\alpha[kf'(t) - k^2 U'(k,t)] + a^2[kf(t) - k^2 U(k,t)] = U'(k,t) \tag{11.8.34}$$

with $U(k,0) = 0$, where the primes denote differentiation with respect to time. Solving Equation 11.8.34,

$$e^{a^2k^2t/(1+\alpha k^2)} U(k,t) = \frac{\alpha k}{1+\alpha k^2} \int_0^t f'(\tau) e^{a^2k^2\tau/(1+\alpha k^2)} d\tau + \frac{a^2 k}{1+\alpha k^2} \int_0^t f(\tau) e^{a^2k^2\tau/(1+\alpha k^2)} d\tau. \tag{11.8.35}$$

[21] See Fetecău, C., and J. Zierep, 2001: On a class of exact solutions of the equations of motion of a second grade fluid. *Acta Mech.*, **150**, 135–138.

Using integration by parts on the second integral in Equation 11.8.35, we find that

$$U(k,t) = \frac{1}{k}\left[f(t) - f(0)e^{-a^2k^2t/(1+\alpha k^2)} - \frac{1}{1+\alpha k^2}\int_0^t f'(\tau)e^{-a^2k^2(t-\tau)/(1+\alpha k^2)}\,d\tau\right]. \tag{11.8.36}$$

Because

$$u(x,t) = \frac{2}{\pi}\int_0^\infty U(k,t)\sin(kx)\,dk, \tag{11.8.37}$$

$$u(x,t) = \frac{2}{\pi}f(t)\int_0^\infty \frac{\sin(kx)}{k}\,dk - \frac{2}{\pi}\int_0^\infty \frac{\sin(kx)}{k}e^{-a^2k^2t/(1+\alpha k^2)}\,dk\left[f(0) + \frac{1}{1+\alpha k^2}\int_0^t f'(\tau)e^{a^2k^2\tau/(1+\alpha k^2)}\,d\tau\right] \tag{11.8.38}$$

$$= f(t) - \frac{2}{\pi}\int_0^\infty \frac{\sin(kx)}{k}e^{-a^2k^2t/(1+\alpha k^2)}dk\left[f(0) + \frac{1}{1+\alpha k^2}\int_0^t f'(\tau)e^{a^2k^2\tau/(1+\alpha k^2)}d\tau\right]. \tag{11.8.39}$$

Problems

For Problems 1–4, find the solution of the heat equation

$$\frac{\partial u}{\partial t} = a^2\frac{\partial^2 u}{\partial x^2}, \qquad -\infty < x < \infty, \quad 0 < t,$$

subject to the stated initial conditions.

1. $u(x,0) = \begin{cases} 1, & |x| < b, \\ 0, & |x| > b. \end{cases}$

2. $u(x,0) = e^{-b|x|}$

3. $u(x,0) = \begin{cases} 0, & -\infty < x < 0, \\ T_0, & 0 < x < b, \\ 0, & b < x < \infty. \end{cases}$

4. $u(x,0) = \delta(x)$

Lovering[22] has applied the solution to Problem 1 to cases involving the cooling of lava.

5. Solve the spherically symmetric equation of diffusion,[23]

$$\frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r}\right), \qquad 0 \le r < \infty, \quad 0 < t,$$

---

[22] Lovering, T. S., 1935: Theory of heat conduction applied to geological problems. *Bull. Geol. Soc. Am.*, **46**, 69–94.

[23] See Shklovskii, I. S., and V. G. Kurt, 1960: Determination of atmospheric density at a height of 430 km by means of the diffusion of sodium vapors. *Am. Rocket Soc. J.*, **30**, 662–667.

with $u(r,0) = u_0(r)$.

Step 1: Assuming $v(r,t) = r\,u(r,t)$, show that the problem can be recast as

$$\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 v}{\partial r^2}, \qquad 0 \le r < \infty, \quad 0 < t,$$

with $v(r,0) = r\,u_0(r)$.

Step 2: Using Equation 11.8.14, show that the general solution is

$$u(r,t) = \frac{1}{2ar\sqrt{\pi t}} \int_0^\infty u_0(\rho) \left\{ \exp\left[-\frac{(r-\rho)^2}{4a^2 t}\right] - \exp\left[-\frac{(r+\rho)^2}{4a^2 t}\right] \right\} \rho\, d\rho.$$

Hint: What is the constraint on Equation 11.8.14 so that the solution remains radially symmetric?

Step 3: For the initial concentration of

$$u_0(r) = \begin{cases} N_0, & 0 \le r < r_0, \\ 0, & r_0 < r, \end{cases}$$

show that

$$u(r,t) = \tfrac{1}{2} N_0 \left[\operatorname{erf}\left(\frac{r_0 - r}{2a\sqrt{t}}\right) + \operatorname{erf}\left(\frac{r_0 + r}{2a\sqrt{t}}\right) + \frac{2a\sqrt{t}}{r\sqrt{\pi}} \left\{ \exp\left[-\frac{(r_0+r)^2}{4a^2 t}\right] - \exp\left[-\frac{(r_0-r)^2}{4a^2 t}\right] \right\} \right],$$

where $\operatorname{erf}(\cdot)$ is the error function.

Further Readings

Bracewell, R. N., 2000: *The Fourier Transform and Its Applications.* McGraw -Hill Book Co., 616 pp. This book presents the theory as well as a wealth of applications.

Körner, T. W., 1988: *Fourier Analyis.* Cambridge University Press, 591 pp. Presents several interesting applications.

Sneddon, I. N., 1995: *Fourier Transforms*. Dover, 542 pp. A wonderful book that illustrates the use of Fourier and Bessel transforms in solving a wealth of problems taken from the sciences and engineering.

Titchmarch, E. C., 1948: *Introduction to the Theory of Fourier Integrals.* Oxford University Press, 391. A source book on the theory of Fourier integrals until 1950.

Chapter 12
The Laplace Transform

The previous chapter introduced the concept of the Fourier integral. If the function is nonzero only when $t > 0$, a similar transform, the *Laplace transform*,[1] exists. It is particularly useful in solving initial-value problems involving linear, constant coefficient, ordinary, and partial differential equations. The present chapter develops the general properties and techniques of Laplace transforms.

12.1 DEFINITION AND ELEMENTARY PROPERTIES

Consider a function $f(t)$ such that $f(t) = 0$ for $t < 0$. Then the *Laplace integral*:

$$\mathcal{L}[f(t)] = F(s) = \int_0^\infty f(t)e^{-st}\,dt \qquad \textbf{(12.1.1)}$$

defines the Laplace transform of $f(t)$, which we shall write $\mathcal{L}[f(t)]$ or $F(s)$. The Laplace transform converts a function of t into a function of the transform variable s.

Not all functions have a Laplace transform because the integral, Equation 12.1.1, may fail to exist. For example, the function may have infinite discontinuities. For this reason, $f(t) = \tan(t)$ does *not* have a Laplace transform. We can avoid this difficulty by requiring that $f(t)$ be *piece-wise continuous*. That is, we can divide a finite range into a finite number of intervals in such a manner that $f(t)$ is continuous inside each interval and approaches finite values as we approach either end of any interval from the interior.

---

[1] The standard reference for Laplace transforms is Doetsch, G., 1950: *Handbuch der Laplace-Transformation. Band 1. Theorie der Laplace-Transformation*. Birkhäuser Verlag, 581 pp.; Doetsch, G., 1955: *Handbuch der Laplace-Transformation. Band 2. Anwendungen der Laplace-Transformation. 1. Abteilung*. Birkhäuser Verlag, 433 pp.; Doetsch, G., 1956: *Handbuch der Laplace-Transformation. Band 3. Anwendungen der Laplace-Transformation. 2. Abteilung*. Birkhäuser Verlag, 298 pp.

Another unacceptable function is $f(t) = 1/t$ because the integral Equation 12.1.1 fails to exist. This leads to the requirement that the product $t^n|f(t)|$ is bounded near $t = 0$ for some number $n < 1$.

Finally, $|f(t)|$ cannot grow too rapidly or it could overwhelm the e^{-st} term. To express this, we introduce the concept of functions of *exponential order*. By exponential order we mean that there exist some constants, M and k, for which $|f(t)| \le Me^{kt}$ for all $t > 0$. Then, the Laplace transform of $f(t)$ exists if s, or just the real part of s, is greater than k.

In summary, the Laplace transform of $f(t)$ exists, for sufficiently large s, provided $f(t)$ satisfies the following conditions:

- $f(t) = 0$ for $t < 0$,
- $f(t)$ is continuous or piece-wise continuous in every interval,
- $t^n|f(t)| < \infty$ as $t \to 0$ for some number n, where $n < 1$,
- $e^{-s_0 t}|f(t)| < \infty$ as $t \to \infty$, for some number s_0. The quantity s_0 is called the *abscissa of convergence*.

• Example 12.1.1

Let us find the Laplace transform of 1, e^{at}, $\sin(at)$, $\cos(at)$, and t^n from the definition of the Laplace transform. From Equation 12.1.1, direct integration yields

$$\mathcal{L}(1) = \int_0^\infty e^{-st}\,dt = -\left.\frac{e^{-st}}{s}\right|_0^\infty = \frac{1}{s}, \qquad s > 0, \tag{12.1.2}$$

$$\mathcal{L}(e^{at}) = \int_0^\infty e^{at}e^{-st}\,dt = \int_0^\infty e^{-(s-a)t}\,dt \tag{12.1.3}$$

$$= -\left.\frac{e^{-(s-a)t}}{s-a}\right|_0^\infty = \frac{1}{s-a}, \qquad s > a, \tag{12.1.4}$$

$$\mathcal{L}[\sin(at)] = \int_0^\infty \sin(at)e^{-st}\,dt = -\left.\frac{e^{-st}}{s^2+a^2}[s\sin(at) + a\cos(at)]\right|_0^\infty \tag{12.1.5}$$

$$= \frac{a}{s^2+a^2}, \qquad s > 0, \tag{12.1.6}$$

$$\mathcal{L}[\cos(at)] = \int_0^\infty \cos(at)e^{-st}\,dt = \left.\frac{e^{-st}}{s^2+a^2}[-s\cos(at) + a\sin(at)]\right|_0^\infty \tag{12.1.7}$$

$$= \frac{s}{s^2+a^2}, \qquad s > 0, \tag{12.1.8}$$

and

$$\mathcal{L}(t^n) = \int_0^\infty t^n e^{-st}\,dt = n!e^{-st}\sum_{m=0}^{n}\left.\frac{t^{n-m}}{(n-m)!s^{m+1}}\right|_0^\infty = \frac{n!}{s^{n+1}}, \quad s > 0, \tag{12.1.9}$$

where n is a positive integer.

MATLAB provides the routine `laplace` to compute the Laplace transform for a given function. For example,

```
>> syms a n s t
>> laplace(1,t,s)
ans =
1/s
>> laplace(exp(a*t),t,s)
ans =
1/(s-a)
>> laplace(sin(a*t),t,s)
ans =
a/(s^2+a^2)
>> laplace(cos(a*t),t,s)
ans =
s/(s^2+a^2)
>> laplace(t^5,t,s)
ans =
120/s^6
```
□

The Laplace transform inherits two important properties from its integral definition. First, the transform of a sum equals the sum of the transforms, or

$$\boxed{\mathcal{L}[c_1 f(t) + c_2 g(t)] = c_1 \mathcal{L}[f(t)] + c_2 \mathcal{L}[g(t)].} \tag{12.1.10}$$

This linearity property holds with complex numbers and functions as well.

• Example 12.1.2

Success with Laplace transforms often rests with the ability to manipulate a given transform into a form that you can invert by inspection. Consider the following examples.

Given $F(s) = 4/s^3$, then

$$F(s) = 2 \times \frac{2}{s^3}, \qquad \text{and} \qquad f(t) = 2t^2 \tag{12.1.11}$$

from Equation 12.1.9.

Given

$$F(s) = \frac{s+2}{s^2+1} = \frac{s}{s^2+1} + \frac{2}{s^2+1}, \tag{12.1.12}$$

then

$$f(t) = \cos(t) + 2\sin(t) \tag{12.1.13}$$

by Equation 12.1.6, Equation 12.1.8, and Equation 12.1.10.

Because

$$F(s) = \frac{1}{s(s-1)} = \frac{1}{s-1} - \frac{1}{s} \tag{12.1.14}$$

by partial fractions, then

$$f(t) = e^t - 1 \tag{12.1.15}$$

by Equation 12.1.2, Equation 12.1.4, and Equation 12.1.10.

MATLAB also provides the routine `ilaplace` to compute the inverse Laplace transform for a given function. For example,

```
>> syms s t
>> ilaplace(4/s^3,s,t)
ans =
2*t^2
>> ilaplace((s+2)/(s^2+1),s,t)
ans =
cos(t)+2*sin(t)
>> ilaplace(1/(s*(s-1)),s,t)
ans =
-1+exp(t)
```

□

The second important property deals with derivatives. Suppose $f(t)$ is continuous and has a piece-wise continuous derivative $f'(t)$. Then

$$\mathcal{L}[f'(t)] = \int_0^\infty f'(t)e^{-st}dt = e^{-st}f(t)\Big|_0^\infty + s\int_0^\infty f(t)e^{-st}dt \tag{12.1.16}$$

by integration by parts. If $f(t)$ is of exponential order, $e^{-st}f(t)$ tends to zero as $t \to \infty$, for large enough s, so that $\mathcal{L}[f'(t)] = sF(s) - f(0)$. Similarly, if $f(t)$ and $f'(t)$ are continuous, $f''(t)$ is piece-wise continuous, and all three functions are of exponential order, then

$$\mathcal{L}[f''(t)] = s\mathcal{L}[f'(t)] - f'(0) = s^2F(s) - sf(0) - f'(0). \tag{12.1.17}$$

In general,

$$\boxed{\mathcal{L}[f^{(n)}(t)] = s^nF(s) - s^{n-1}f(0) - \cdots - sf^{(n-2)}(0) - f^{(n-1)}(0)} \tag{12.1.18}$$

on the assumption that $f(t)$ and its first $n-1$ derivatives are continuous, $f^{(n)}(t)$ is piece-wise continuous, and all are of exponential order so that the Laplace transform exists.

The converse of Equation 12.1.18 is also of some importance. If

$$u(t) = \int_0^t f(\tau)\,d\tau, \tag{12.1.19}$$

then

$$\mathcal{L}[u(t)] = \int_0^\infty e^{-st}\left[\int_0^t f(\tau)\,d\tau\right]dt = -\left.\frac{e^{-st}}{s}\int_0^t f(\tau)\,d\tau\right|_0^\infty + \frac{1}{s}\int_0^\infty f(t)e^{-st}dt, \tag{12.1.20}$$

and

$$\boxed{\mathcal{L}\left[\int_0^t f(\tau)\,d\tau\right] = F(s)/s,} \tag{12.1.21}$$

where $u(0) = 0$.

Problems

Using the definition of the Laplace transform, find the Laplace transform of the following functions. For Problems 1–4, check your answers using MATLAB.

1. $f(t) = \cosh(at)$

2. $f(t) = \cos^2(at)$

3. $f(t) = (t+1)^2$

4. $f(t) = (t+1)e^{-at}$

5. $f(t) = \begin{cases} e^t, & 0 < t < 2 \\ 0, & 2 < t \end{cases}$

6. $f(t) = \begin{cases} \sin(t), & 0 \le t \le \pi \\ 0, & \pi \le t \end{cases}$

Using your knowledge of the transform for 1, e^{at}, $\sin(at)$, $\cos(at)$, and t^n, find the Laplace transform of

7. $f(t) = 2\sin(t) - \cos(2t) + \cos(3) - t$

8. $f(t) = t - 2 + e^{-5t} - \sin(5t) + \cos(2)$.

Find the inverse of the following transforms. Verify your result using MATLAB.

9. $F(s) = 1/(s+3)$

10. $F(s) = 1/s^4$

11. $F(s) = 1/(s^2+9)$

12. $F(s) = (2s+3)/(s^2+9)$

13. $F(s) = 2/(s^2+1) - 15/s^3 + 2/(s+1) - 6s/(s^2+4)$

14. $F(s) = 3/s + 15/s^3 + (s+5)/(s^2+1) - 6/(s-2)$.

15. Verify the derivative rule for Laplace transforms using the function $f(t) = \sin(at)$.

16. Show that $\mathcal{L}[f(at)] = F(s/a)/a$, where $F(s) = \mathcal{L}[f(t)]$.

17. Using the trigonometric identity $\sin^2(x) = [1 - \cos(2x)]/2$, find the Laplace transform of $f(t) = \sin^2[\pi t/(2T)]$.

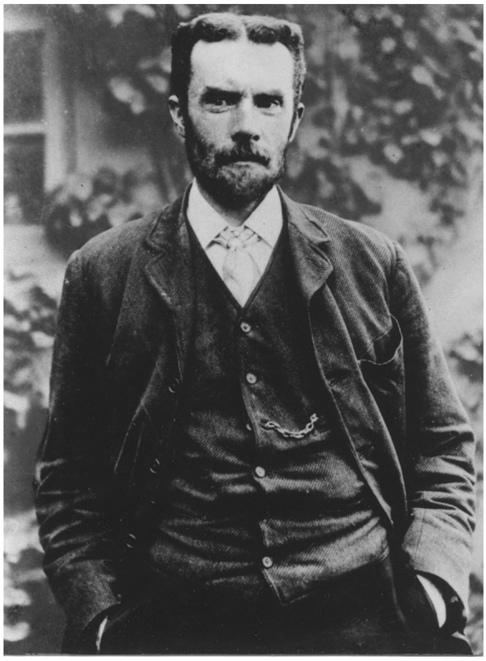
12.2 THE HEAVISIDE STEP AND DIRAC DELTA FUNCTIONS

Change can occur abruptly. We throw a switch and electricity suddenly flows. In this section we introduce two functions, the Heaviside step and Dirac delta, that will give us the ability to construct complicated discontinuous functions to express these changes.

Heaviside step function

We define the *Heaviside step function* as

$$H(t-a) = \begin{cases} 1, & t > a, \\ 0, & t < a, \end{cases} \tag{12.2.1}$$



Largely a self-educated man, Oliver Heaviside (1850–1925) lived the life of a recluse. It was during his studies of the implications of Maxwell's theory of electricity and magnetism that he re-invented Laplace transforms. Initially rejected, it would require the work of Bromwich to justify its use. (Portrait courtesy of the Institution of Engineering and Technology Archives.)

where $a \geq 0$. From this definition,

$$\mathcal{L}[H(t-a)] = \int_a^\infty e^{-st}\,dt = \frac{e^{-as}}{s}, \qquad s > 0. \tag{12.2.2}$$

Note that this transform is identical to that for $f(t) = 1$ if $a = 0$. This should not surprise us. As pointed out earlier, the function $f(t)$ is zero for all $t < 0$ by definition. Thus, when dealing with Laplace transforms, $f(t) = 1$ and $H(t)$ are identical. Generally we will take 1 rather than $H(t)$ as the inverse of $1/s$. The Heaviside step function is essentially a bookkeeping device that gives us the ability to "switch on" and "switch off" a given function. For example, if we want a function $f(t)$ to become nonzero at time $t = a$, we represent this process by the product $f(t)H(t-a)$. On the other hand, if we only want the function to be "turned on" when $a < t < b$, the desired expression is then $f(t)[H(t-a) - H(t-b)]$. For $t < a$, both step functions in the brackets have the value of zero. For $a < t < b$, the first step function has the value of unity and the second step function has the value of zero, so that we have $f(t)$. For $t > b$, both step functions equal unity so that their difference is zero.

• Example 12.2.1

Quite often we need to express the graphical representation of a function by a mathematical equation. We can conveniently do this through the use of step functions in a two-step procedure. The following example illustrates this procedure.

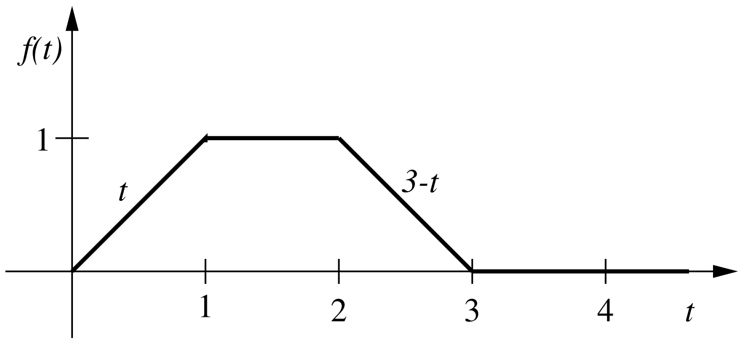


Figure 12.2.1: Graphical representation of Equation 12.2.5.

Consider Figure 12.2.1. We would like to express this graph in terms of Heaviside step functions. We begin by introducing step functions at each point where there is a kink (discontinuity in the first derivative) or jump in the graph - in the present case at $t = 0$, $t = 1$, $t = 2$, and $t = 3$. These are the points of abrupt change. Thus,

$$f(t) = a_0(t)H(t) + a_1(t)H(t-1) + a_2(t)H(t-2) + a_3(t)H(t-3), \tag{12.2.3}$$

where the coefficients $a_0(t), a_1(t), \ldots$ are yet to be determined. Proceeding from left to right in Figure 12.2.1, the coefficient of each step function equals the mathematical expression that we want after the kink or jump minus the expression before the kink or jump. As each Heaviside turns on, we need to add in the new t behavior and subtract out the old t behavior. Thus, in the present example,

$$f(t) = (t-0)H(t) + (1-t)H(t-1) + [(3-t)-1]H(t-2) + [0-(3-t)]H(t-3) \tag{12.2.4}$$

or

$$f(t) = tH(t) - (t-1)H(t-1) - (t-2)H(t-2) + (t-3)H(t-3). \tag{12.2.5}$$

We can easily find the Laplace transform of Equation 12.2.5 by the "second shifting" theorem introduced in the next section. □

• Example 12.2.2

Laplace transforms are particularly useful in solving initial-value problems involving linear, constant coefficient, ordinary differential equations where the nonhomogeneous term is discontinuous. As we shall show in the next section, we must first rewrite the nonhomogeneous term using the Heaviside step function before we can use Laplace transforms. For example, given the nonhomogeneous ordinary differential equation:

$$y'' + 3y' + 2y = \begin{cases} t, & 0 < t < 1 \\ 0, & 1 < t, \end{cases} \tag{12.2.6}$$

we can rewrite the right side of Equation 12.2.6 as

$$y'' + 3y' + 2y = t - tH(t-1) = t - (t-1)H(t-1) - H(t-1). \tag{12.2.7}$$

In Section 12.8 we will show how to solve this type of ordinary differential equation using Laplace transforms. □

The Laplace Transforms of Some Commonly Encountered Functions

	$f(t),\ t \geq 0$	$F(s)$
1.	1	$\frac{1}{s}$
2.	e^{-at}	$\frac{1}{s+a}$
3.	$\frac{1}{a}\left(1-e^{-at}\right)$	$\frac{1}{s(s+a)}$
4.	$\frac{1}{a-b}(e^{-bt}-e^{-at})$	$\frac{1}{(s+a)(s+b)}$
5.	$\frac{1}{b-a}(be^{-bt}-ae^{-at})$	$\frac{s}{(s+a)(s+b)}$
6.	$\sin(at)$	$\frac{a}{s^2+a^2}$
7.	$\cos(at)$	$\frac{s}{s^2+a^2}$
8.	$\sinh(at)$	$\frac{a}{s^2-a^2}$
9.	$\cosh(at)$	$\frac{s}{s^2-a^2}$
10.	$t\sin(at)$	$\frac{2as}{(s^2+a^2)^2}$
11.	$1-\cos(at)$	$\frac{a^2}{s(s^2+a^2)}$
12.	$at-\sin(at)$	$\frac{a^3}{s^2(s^2+a^2)}$
13.	$t\cos(at)$	$\frac{s^2-a^2}{(s^2+a^2)^2}$
14.	$\sin(at)-at\cos(at)$	$\frac{2a^3}{(s^2+a^2)^2}$
15.	$t\sinh(at)$	$\frac{2as}{(s^2-a^2)^2}$
16.	$t\cosh(at)$	$\frac{s^2+a^2}{(s^2-a^2)^2}$
17.	$at\cosh(at)-\sinh(at)$	$\frac{2a^3}{(s^2-a^2)^2}$
18.	$e^{-bt}\sin(at)$	$\frac{a}{(s+b)^2+a^2}$
19.	$e^{-bt}\cos(at)$	$\frac{s+b}{(s+b)^2+a^2}$
20.	$(1+a^2t^2)\sin(at)-at\cos(at)$	$\frac{8a^3s^2}{(s^2+a^2)^3}$

The Laplace Transforms of Some Commonly Encountered Functions (Continued)

	$\mathbf{f(t),\ t \geq 0}$	$\mathbf{F(s)}$
21.	$\sin(at)\cosh(at) - \cos(at)\sinh(at)$	$\frac{4a^3}{s^4 + 4a^4}$
22.	$\sin(at)\sinh(at)$	$\frac{2a^2 s}{s^4 + 4a^4}$
23.	$\sinh(at) - \sin(at)$	$\frac{2a^3}{s^4 - a^4}$
24.	$\cosh(at) - \cos(at)$	$\frac{2a^2 s}{s^4 - a^4}$
25.	$\frac{a\sin(at) - b\sin(bt)}{a^2 - b^2}, a^2 \neq b^2$	$\frac{s^2}{(s^2 + a^2)(s^2 + b^2)}$
26.	$\frac{b\sin(at) - a\sin(bt)}{ab(b^2 - a^2)}, a^2 \neq b^2$	$\frac{1}{(s^2 + a^2)(s^2 + b^2)}$
27.	$\frac{\cos(at) - \cos(bt)}{b^2 - a^2}, a^2 \neq b^2$	$\frac{s}{(s^2 + a^2)(s^2 + b^2)}$
28.	$t^n, n \geq 0$	$\frac{n!}{s^{n+1}}$
29.	$\frac{t^{n-1}e^{-at}}{(n-1)!}, n > 0$	$\frac{1}{(s+a)^n}$
30.	$\frac{(n-1) - at}{(n-1)!}t^{n-2}e^{-at}, n > 1$	$\frac{s}{(s+a)^n}$
31.	$t^n e^{-at}, n \geq 0$	$\frac{n!}{(s+a)^{n+1}}$
32.	$\frac{2^n t^{n-(1/2)}}{1 \cdot 3 \cdot 5 \cdots (2n-1)\sqrt{\pi}}, n \geq 1$	$s^{-[n+(1/2)]}$
33.	$J_0(at)$	$\frac{1}{\sqrt{s^2 + a^2}}$
34.	$I_0(at)$	$\frac{1}{\sqrt{s^2 - a^2}}$
35.	$\frac{1}{\sqrt{a}}\ \mathrm{erf}(\sqrt{at}\,)$	$\frac{1}{s\sqrt{s+a}}$
36.	$\frac{1}{\sqrt{\pi t}}e^{-at} + \sqrt{a}\ \mathrm{erf}(\sqrt{at}\,)$	$\frac{\sqrt{s+a}}{s}$
37.	$\frac{1}{\sqrt{\pi t}} - ae^{a^2 t}\mathrm{erfc}(a\sqrt{t}\,)$	$\frac{1}{a + \sqrt{s}}$
38.	$e^{at}\mathrm{erfc}(\sqrt{at}\,)$	$\frac{1}{s + \sqrt{as}}$
39.	$\frac{1}{2\sqrt{\pi t^3}}\left(e^{bt} - e^{at}\right)$	$\sqrt{s-a} - \sqrt{s-b}$

The Laplace Transforms of Some Commonly Encountered Functions (Continued)

	$\mathbf{f(t),\ t \geq 0}$	$\mathbf{F(s)}$
40.	$\dfrac{1}{\sqrt{\pi t}} + ae^{a^2 t}\operatorname{erf}(a\sqrt{t}\,)$	$\dfrac{\sqrt{s}}{s-a^2}$
41.	$\dfrac{1}{\sqrt{\pi t}}e^{at}(1+2at)$	$\dfrac{s}{(s-a)\sqrt{s-a}}$
42.	$\dfrac{1}{a}e^{a^2 t}\operatorname{erf}(a\sqrt{t}\,)$	$\dfrac{1}{(s-a^2)\sqrt{s}}$
43.	$\sqrt{\dfrac{a}{\pi t^3}}e^{-a/t}, a > 0$	$e^{-2\sqrt{as}}$
44.	$\dfrac{1}{\sqrt{\pi t}}e^{-a/t}, a \geq 0$	$\dfrac{1}{\sqrt{s}}e^{-2\sqrt{as}}$
45.	$\operatorname{erfc}\left(\sqrt{\dfrac{a}{t}}\right), a \geq 0$	$\dfrac{1}{s}e^{-2\sqrt{as}}$
46.	$2\sqrt{\dfrac{t}{\pi}}\exp\left(-\dfrac{a^2}{4t}\right) - a\operatorname{erfc}\left(\dfrac{a}{2\sqrt{t}}\right),\ a \geq 0$	$\dfrac{e^{-a\sqrt{s}}}{s\sqrt{s}}$
47.	$-e^{b^2 t+ab}\operatorname{erfc}\left(b\sqrt{t}+\frac{a}{2\sqrt{t}}\right) + \operatorname{erfc}\left(\frac{a}{2\sqrt{t}}\right), a \geq 0$	$\dfrac{be^{-a\sqrt{s}}}{s(b+\sqrt{s}\,)}$
48.	$e^{ab}e^{b^2 t}\operatorname{erfc}\left(b\sqrt{t}+\dfrac{a}{2\sqrt{t}}\right), a \geq 0$	$\dfrac{e^{-a\sqrt{s}}}{\sqrt{s}\,(b+\sqrt{s}\,)}$

Notes:

$$\text{Error function: } \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-y^2}\,dy$$

$$\text{Complementary error function: } \operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$$

Dirac delta function

The second special function is the *Dirac delta function* or *impulse function*. We define it by

$$\delta(t-a) = \begin{cases} \infty, & t = a, \\ 0, & t \neq a, \end{cases} \qquad \int_0^\infty \delta(t-a)\,dt = 1, \tag{12.2.8}$$

where $a \geq 0$.

A popular way of visualizing the delta function is as a very narrow rectangular pulse:

$$\delta(t-a) = \lim_{\epsilon \to 0} \begin{cases} 1/\epsilon, & 0 < |t-a| < \epsilon/2, \\ 0, & |t-a| > \epsilon/2, \end{cases} \tag{12.2.9}$$

where $\epsilon > 0$ is some small number and $a > 0$. See Figure 12.2.2. This pulse has a width ϵ, height $1/\epsilon$, and its center at $t = a$ so that its area is unity. Now, as this pulse shrinks in width ($\epsilon \to 0$), its height increases so that it remains centered at $t = a$ and its area equals unity. If we continue this process, always keeping the area unity and the pulse

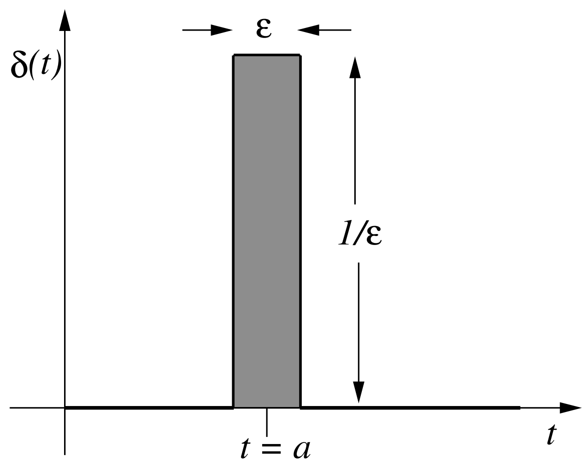


Figure 12.2.2: The Dirac delta function.

symmetric about $t = a$, eventually we obtain an extremely narrow, very large amplitude pulse at $t = a$. If we proceed to the limit, where the width approaches zero and the height approaches infinity (but still with unit area), we obtain the delta function $\delta(t-a)$.

The delta function was introduced earlier during our study of Fourier transforms. So what is the difference between the delta function introduced then and the delta function now? Simply put, the delta function can now only be used on the interval $[0, \infty)$. Outside of that, we shall use it very much as we did with Fourier transforms.

Using Equation 12.2.9, the Laplace transform of the delta function is

$$\mathcal{L}[\delta(t-a)] = \int_0^\infty \delta(t-a)e^{-st}dt = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_{a-\epsilon/2}^{a+\epsilon/2} e^{-st}dt \tag{12.2.10}$$

$$= \lim_{\epsilon\to 0} \quad \frac{1}{\epsilon s}\left(e^{-as+\epsilon s/2} - e^{-as-\epsilon s/2}\right) \tag{12.2.11}$$

$$= \lim_{\epsilon\to 0} \quad \frac{e^{-as}}{\epsilon s}\left(1 + \frac{\epsilon s}{2} + \frac{\epsilon^2 s^2}{8} + \cdots - 1 + \frac{\epsilon s}{2} - \frac{\epsilon^2 s^2}{8} + \cdots\right) \tag{12.2.12}$$

$$= e^{-as}. \tag{12.2.13}$$

In the special case when $a = 0$, $\mathcal{L}[\delta(t)] = 1$, a property that we will use in Section 12.9. Note that this is exactly the result that we obtained for the Fourier transform of the delta function.

If we integrate the impulse function,

$$\int_0^t \delta(\tau - a)\, d\tau = \begin{cases} 0, & t < a, \\ 1, & t > a, \end{cases} \tag{12.2.14}$$

according to whether the impulse does or does not come within the range of integration. This integral gives a result that is precisely the definition of the Heaviside step function, so that we can rewrite Equation 12.2.14:

$$\int_0^t \delta(\tau - a)\, d\tau = H(t-a). \tag{12.2.15}$$

Consequently, the delta function behaves like the derivative of the step function, or

$$\frac{d}{dt}\left[H(t-a)\right] = \delta(t-a). \tag{12.2.16}$$

Because the conventional derivative does not exist at a point of discontinuity, we can only make sense of Equation 12.2.16 if we extend the definition of the derivative. Here we extended the definition formally, but a richer and deeper understanding arises from the theory of generalized functions.[2]

• Example 12.2.3

Let us find the (generalized) derivative of

$$f(t) = 3t^2\left[H(t) - H(t-1)\right]. \tag{12.2.17}$$

Proceeding formally,

$$f'(t) = 6t\left[H(t) - H(t-1)\right] + 3t^2\left[\delta(t) - \delta(t-1)\right] \tag{12.2.18}$$
$$= 6t\left[H(t) - H(t-1)\right] + 0 - 3\delta(t-1) \tag{12.2.19}$$
$$= 6t\left[H(t) - H(t-1)\right] - 3\delta(t-1), \tag{12.2.20}$$

because $f(t)\delta(t-t_0) = f(t_0)\delta(t-t_0)$. □

• Example 12.2.4

MATLAB also includes the step and Dirac delta functions among its intrinsic functions. There are two types of step functions. In symbolic calculations, the function is `Heaviside` while `stepfunction` is used in numerical calculations. For example, the Laplace transform of Equation 12.2.5 is

```
>>syms s,t
>>laplace('t-(t-1)*Heaviside(t-1)-(t-2)*Heaviside(t-2)'...
    '+(t-3)*Heaviside(t-3)',t,s)
ans =
1/s^2-exp(-s)/s^2-exp(-2*s)/s^2+exp(-3*s)/s^2
```

In a similar manner, the symbolic function for the Dirac delta function is `Dirac`. Therefore, the Laplace transform of $(t-1)\delta(t-2)$ is

```
>>syms s,t
>>laplace('(t-1)*Dirac(t-2)',t,s)
ans =
exp(-2*s)
```

[2] The generalization of the definition of a function so that it can express in a mathematically correct form such idealized concepts as the density of a material point, a point charge or point dipole, the space charge of a simple or double layer, the intensity of an instantaneous source, etc.

Problems

Sketch the following functions and express them in terms of the Heaviside step functions:

1. $f(t) = \begin{cases} 0, & 0 \le t \le 2 \\ t-2, & 2 \le t < 3 \\ 0, & 3 < t \end{cases}$

2. $f(t) = \begin{cases} 0, & 0 < t < a \\ 1, & a < t < 2a \\ -1, & 2a < t < 3a \\ 0, & 3a < t \end{cases}$

Rewrite the following nonhomogeneous ordinary differential equations using the Heaviside step functions:

3. $y'' + 3y' + 2y = \begin{cases} 0, & 0 < t < 1 \\ 1, & 1 < t \end{cases}$

4. $y'' + 4y = \begin{cases} 0, & 0 < t < 4 \\ 3, & 4 < t \end{cases}$

5. $y'' + 4y' + 4y = \begin{cases} 0, & 0 < t < 2 \\ t, & 2 < t \end{cases}$

6. $y'' + 3y' + 2y = \begin{cases} 0, & 0 < t < 1 \\ e^t, & 1 < t \end{cases}$

7. $y'' - 3y' + 2y = \begin{cases} 0, & 0 < t < 2 \\ e^{-t}, & 2 < t \end{cases}$

8. $y'' - 3y' + 2y = \begin{cases} 0, & 0 < t < 1 \\ t^2, & 1 < t \end{cases}$

9. $y'' + y = \begin{cases} \sin(t), & 0 \le t \le \pi \\ 0, & \pi \le t \end{cases}$

10. $y'' + 3y' + 2y = \begin{cases} t, & 0 \le t \le a \\ ae^{-(t-a)}, & a \le t \end{cases}$

12.3 SOME USEFUL THEOREMS

Although at first sight there would appear to be a bewildering number of transforms to either memorize or tabulate, there are several useful theorems, that can extend the applicability of a given transform.

First shifting theorem

Consider the transform of the function $e^{-at}f(t)$, where a is any real number. Then, by definition,

$$\mathcal{L}\left[e^{-at}f(t)\right] = \int_0^\infty e^{-st}e^{-at}f(t)\,dt = \int_0^\infty e^{-(s+a)t}f(t)\,dt, \tag{12.3.1}$$

or

$$\mathcal{L}\left[e^{-at}f(t)\right] = F(s+a). \tag{12.3.2}$$

That is, if $F(s)$ is the transform of $f(t)$ and a is a constant, then $F(s+a)$ is the transform of $e^{-at}f(t)$.

• Example 12.3.1

Let us find the Laplace transform of $f(t) = e^{-at}\sin(bt)$. Because the Laplace transform of $\sin(bt)$ is $b/(s^2+b^2)$,

$$\mathcal{L}\left[e^{-at}\sin(bt)\right] = \frac{b}{(s+a)^2+b^2}, \tag{12.3.3}$$

where we simply replaced s by $s + a$ in the transform for $\sin(bt)$. □

• Example 12.3.2

Let us find the inverse of the Laplace transform

$$F(s) = \frac{s+2}{s^2+6s+1}. \tag{12.3.4}$$

Rearranging terms,

$$F(s) = \frac{s+2}{s^2+6s+1} = \frac{s+2}{(s+3)^2-8} = \frac{s+3}{(s+3)^2-8} - \frac{1}{2\sqrt{2}}\frac{2\sqrt{2}}{(s+3)^2-8}. \tag{12.3.5}$$

Immediately, from the first shifting theorem,

$$f(t) = e^{-3t}\cosh(2\sqrt{2}t) - \frac{e^{-3t}}{2\sqrt{2}}\sinh(2\sqrt{2}t). \tag{12.3.6}$$

□

Second shifting theorem

The *second shifting theorem* states that if $F(s)$ is the transform of $f(t)$, then $e^{-bs}F(s)$ is the transform of $f(t-b)H(t-b)$, where b is real and positive. To show this, consider the Laplace transform of $f(t-b)H(t-b)$. Then, from the definition,

$$\mathcal{L}[f(t-b)H(t-b)] = \int_0^\infty f(t-b)H(t-b)e^{-st}\,dt \tag{12.3.7}$$

$$= \int_b^\infty f(t-b)e^{-st}\,dt = \int_0^\infty e^{-bs}e^{-sx}f(x)\,dx \tag{12.3.8}$$

$$= e^{-bs}\int_0^\infty e^{-sx}f(x)\,dx, \tag{12.3.9}$$

or

$$\mathcal{L}[f(t-b)H(t-b)] = e^{-bs}F(s), \tag{12.3.10}$$

where we set $x = t - b$. This theorem is of fundamental importance because it allows us to write down the transforms for "delayed" time functions. That is, functions that "turn on" b units after the initial time.

• Example 12.3.3

Let us find the inverse of the transform $(1 - e^{-s})/s$. Since

$$\frac{1-e^{-s}}{s} = \frac{1}{s} - \frac{e^{-s}}{s}, \tag{12.3.11}$$

$$\mathcal{L}^{-1}\left(\frac{1}{s} - \frac{e^{-s}}{s}\right) = \mathcal{L}^{-1}\left(\frac{1}{s}\right) - \mathcal{L}^{-1}\left(\frac{e^{-s}}{s}\right) = H(t) - H(t-1), \tag{12.3.12}$$

because $\mathcal{L}^{-1}(1/s) = f(t) = 1$, and $f(t-1) = 1$. □

• **Example 12.3.4**

Let us find the Laplace transform of $f(t) = (t^2 - 1)H(t-1)$.
We begin by noting that

$$
\begin{aligned}
(t^2-1)H(t-1) &= [(t-1+1)^2-1]H(t-1) && \textbf{(12.3.13)}\\
&= [(t-1)^2+2(t-1)]H(t-1) && \textbf{(12.3.14)}\\
&= (t-1)^2H(t-1)+2(t-1)H(t-1). && \textbf{(12.3.15)}
\end{aligned}
$$

A direct application of the second shifting theorem then leads to

$$\mathcal{L}[(t^2-1)H(t-1)] = \frac{2e^{-s}}{s^3} + \frac{2e^{-s}}{s^2}. \qquad \textbf{(12.3.16)}$$

□

• **Example 12.3.5**

In Example 12.2.2 we discussed the use of Laplace transforms in solving ordinary differential equations. One further step along the road consists of finding $Y(s) = \mathcal{L}[y(t)]$. Now that we have the second shifting theorem, let us do this.

Continuing Example 12.2.2 with $y(0) = 0$ and $y'(0) = 1$, let us take the Laplace transform of Equation 12.2.7. Employing the second shifting theorem and Equation 12.1.18, we find that

$$s^2Y(s) - sy(0) - y'(0) + 3sY(s) - 3y(0) + 2Y(s) = \frac{1}{s^2} - \frac{e^{-s}}{s^2} - \frac{e^{-s}}{s}. \qquad \textbf{(12.3.17)}$$

Substituting in the initial conditions and solving for $Y(s)$, we finally obtain

$$Y(s) = \frac{1}{(s+2)(s+1)} + \frac{1}{s^2(s+2)(s+1)} + \frac{e^{-s}}{s^2(s+2)(s+1)} + \frac{e^{-s}}{s(s+2)(s+1)}. \qquad \textbf{(12.3.18)}$$

□

Laplace transform of $t^n f(t)$

In addition to the shifting theorems, there are two other particularly useful theorems that involve the derivative and integral of the transform $F(s)$. For example, if we write

$$F(s) = \mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}\,dt \qquad \textbf{(12.3.19)}$$

and differentiate with respect to s, then

$$F'(s) = \int_0^\infty -tf(t)e^{-st}\,dt = -\mathcal{L}[tf(t)]. \qquad \textbf{(12.3.20)}$$

In general, we have that

$$F^{(n)}(s) = (-1)^n\mathcal{L}[t^nf(t)]. \qquad \textbf{(12.3.21)}$$

Laplace transform of $f(t)/t$

Consider the following integration of the Laplace transform $F(s)$:

$$\int_s^\infty F(z)\,dz = \int_s^\infty \left[\int_0^\infty f(t)e^{-zt}dt\right] dz. \tag{12.3.22}$$

Upon interchanging the order of integration, we find that

$$\int_s^\infty F(z)\,dz = \int_0^\infty f(t)\left[\int_s^\infty e^{-zt}dz\right] dt = -\int_0^\infty f(t)\left.\frac{e^{-zt}}{t}\right|_s^\infty dt = \int_0^\infty \frac{f(t)}{t}e^{-st}dt. \tag{12.3.23}$$

Therefore,

$$\int_s^\infty F(z)\,dz = \mathcal{L}\left[\frac{f(t)}{t}\right]. \tag{12.3.24}$$

• **Example 12.3.6**

Let us find the transform of $t\sin(at)$. From Equation 12.3.20,

$$\mathcal{L}[t\sin(at)] = -\frac{d}{ds}\Big\{\mathcal{L}[\sin(at)]\Big\} = -\frac{d}{ds}\left[\frac{a}{s^2+a^2}\right] = \frac{2as}{(s^2+a^2)^2}. \tag{12.3.25}$$

□

• **Example 12.3.7**

Let us find the transform of $[1-\cos(at)]/t$. To solve this problem, we apply Equation 12.3.24 and find that

$$\mathcal{L}\left[\frac{1-\cos(at)}{t}\right] = \int_s^\infty \mathcal{L}[1-\cos(at)]\Big|_{s=z} dz = \int_s^\infty \left(\frac{1}{z} - \frac{z}{z^2+a^2}\right) dz \tag{12.3.26}$$

$$= \ln(z) - \tfrac{1}{2}\ln(z^2+a^2)\Big|_s^\infty = \ln\left(\frac{z}{\sqrt{z^2+a^2}}\right)\Bigg|_s^\infty \tag{12.3.27}$$

$$= \ln(1) - \ln\left(\frac{s}{\sqrt{s^2+a^2}}\right) = -\ln\left(\frac{s}{\sqrt{s^2+a^2}}\right). \tag{12.3.28}$$

□

Initial-value theorem

Let $f(t)$ and $f'(t)$ possess Laplace transforms. Then, from the definition of the Laplace transform,

$$\int_0^\infty f'(t)e^{-st}\,dt = sF(s) - f(0). \tag{12.3.29}$$

Because s is a parameter in Equation 12.3.29 and the existence of the integral is implied by the derivative rule, we can let $s\to\infty$ before we integrate. In that case, the left side of Equation 12.3.29 vanishes to zero, which leads to

$$\lim_{s\to\infty} sF(s) = f(0). \tag{12.3.30}$$

This is the *initial-value theorem.*

• Example 12.3.8

Let us verify the initial-value theorem using $f(t) = e^{3t}$. Because $F(s) = 1/(s-3)$, $\lim_{s\to\infty} s/(s-3) = 1$. This agrees with $f(0) = 1$.

In the common case when the Laplace transform is ratio to two polynomials, we can use MATLAB to find the initial value. This consists of two steps. First, we construct $sF(s)$ by creating vectors that describe the numerator and denominator of $sF(s)$ and then evaluate the numerator and denominator using very large values of s. For example, in the previous example,

```
>>num = [1 0];
>>den = [1 -3];
>>initialvalue = polyval(num,1e20) / polyval(den,1e20)
initialvalue =
     1
```

□

Final-value theorem

Let $f(t)$ and $f'(t)$ possess Laplace transforms. Then, in the limit of $s \to 0$, Equation 12.3.29 becomes

$$\int_0^\infty f'(t)\,dt = \lim_{t\to\infty}\int_0^t f'(\tau)\,d\tau = \lim_{t\to\infty} f(t) - f(0) = \lim_{s\to 0} sF(s) - f(0). \tag{12.3.31}$$

Because $f(0)$ is not a function of t or s, the quantity $f(0)$ cancels from Equation 12.3.31, leaving

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s). \tag{12.3.32}$$

Equation 12.3.32 is the *final-value theorem*. It should be noted that this theorem assumes that $\lim_{t\to\infty} f(t)$ exists. For example, it does not apply to sinusoidal functions. Thus, we must restrict ourselves to Laplace transforms that have singularities in the left half of the s-plane unless they occur at the origin.

In the case when $f(t)$ is a periodic function with a period T, Gluskin[3] showed that

$$\lim_{s\to 0} sF(s) = \frac{1}{T}\int_0^T f(t)\,dt. \tag{12.3.33}$$

• Example 12.3.9

Let us verify the final-value theorem using $f(t) = t$. Because $F(s) = 1/s^2$,

$$\lim_{s\to 0} sF(s) = \lim_{s\to 0} 1/s = \infty. \tag{12.3.34}$$

The limit of $f(t)$ as $t \to \infty$ is also undefined.

[3] Gluskin, E., 2003: Let us teach this generalization of the final-value theorem. *Eur. J. Phys.*, **24**, 591–597.

Just as we can use MATLAB to find the initial value of a Laplace transform in the case when $F(s)$ is a ratio of two polynomials, we can do the same here for the final value. Again we define vectors `num` and `den` that give $sF(s)$ and then evaluate them at $s = 0$. Using the previous example, the MATLAB commands are:

```
>>num = [0 1 0];
>>den = [1 0 0];
>>finalvalue = polyval(num,0) / polyval(den,0)
Warning:  Divide by zero.
finalvalue =
   NaN
```

This agrees with the result from a hand calculation and shows what happens when the denominator has a zero. □

• Example 12.3.10

Looking ahead, we will shortly need to find the Laplace transform of $y(t)$, which is defined by a differential equation. For example, we will want $Y(s)$ where $y(t)$ is governed by

$$y'' + 2y' + 2y = \cos(t) + \delta(t - \pi/2), \qquad y(0) = y'(0) = 0. \tag{12.3.35}$$

Applying Laplace transforms to both sides of Equation 12.3.35, we have that

$$\mathcal{L}(y'') + 2\mathcal{L}(y') + 2\mathcal{L}(y) = \mathcal{L}[\cos(t)] + \mathcal{L}[\delta(t - \pi/2)], \tag{12.3.36}$$

or

$$s^2Y(s) - sy(0) - y'(0) + 2sY(s) - 2y(0) + 2Y(s) = \frac{s}{s^2+1} + e^{-s\pi/2}. \tag{12.3.37}$$

Substituting for $y(0)$ and $y'(0)$ and solving for $Y(s)$, we find that

$$Y(s) = \frac{s}{(s^2+1)(s^2+2s+2)} + \frac{e^{-s\pi/2}}{s^2+2s+2}. \tag{12.3.38}$$

Presently this is as far as we can go.

How would we use MATLAB to find $Y(s)$? The following MATLAB script shows you how:

```
clear
% define symbolic variables
syms pi s t Y
% take Laplace transform of left side of differential equation
LHS = laplace(diff(diff(sym('y(t)')))+2*diff(sym('y(t)'))...
    +2*sym('y(t)'));
% take Laplace transform of right side of differential equation
RHS = laplace(cos(t)+'Dirac(t-pi/2)',t,s);
% set Y for Laplace transform of y
%     and introduce initial conditions
newLHS = subs(LHS,'laplace(y(t),t,s)','y(0)','D(y)(0)',Y,0,0);
% solve for Y
Y = solve(newLHS-RHS,Y)
```

It yields

```
Y =
(s+exp(-1/2*pi*s)*s^2+exp(-1/2*pi*s))/(s^4+3*s^2+2*s^3+2*s+2)
```

Problems

Find the Laplace transform of the following functions and then check your work using MATLAB.

1. $f(t) = e^{-t}\sin(2t)$

2. $f(t) = e^{-2t}\cos(2t)$

3. $f(t) = t^2 H(t-1)$

4. $f(t) = e^{2t} H(t-3)$

5. $f(t) = te^t + \sin(3t)e^t + \cos(5t)e^{2t}$

6. $f(t) = t^4 e^{-2t} + \sin(3t)e^t + \cos(4t)e^{2t}$

7. $f(t) = t^2 e^{-t} + \sin(2t)e^t + \cos(3t)e^{-3t}$

8. $f(t) = t^2 H(t-1) + e^t H(t-2)$

9. $f(t) = (t^2+2)H(t-1) + H(t-2)$

10. $f(t) = (t+1)^2 H(t-1) + e^t H(t-2)$

11. $f(t) = \begin{cases} \sin(t), & 0 \le t \le \pi \\ 0, & \pi \le t \end{cases}$

12. $f(t) = \begin{cases} t, & 0 \le t \le 2 \\ 2, & 2 \le t \end{cases}$

13. $f(t) = te^{-3t}\sin(2t)$

Find the inverse of the following Laplace transforms by hand and using MATLAB:

14. $F(s) = \dfrac{1}{(s+2)^4}$

15. $F(s) = \dfrac{s}{(s+2)^4}$

16. $F(s) = \dfrac{s}{s^2+2s+2}$

17. $F(s) = \dfrac{s+3}{s^2+2s+2}$

18. $F(s) = \dfrac{s}{(s+1)^3} + \dfrac{s+1}{s^2+2s+2}$

19. $F(s) = \dfrac{s}{(s+2)^2} + \dfrac{s+2}{s^2+2s+2}$

20. $F(s) = \dfrac{s}{(s+2)^3} + \dfrac{s+4}{s^2+4s+5}$

21. $F(s) = \dfrac{e^{-3s}}{s-1}$

22. $F(s) = \dfrac{e^{-2s}}{(s+1)^2}$

23. $F(s) = \dfrac{s\,e^{-s}}{s^2+2s+2}$

24. $F(s) = \dfrac{e^{-4s}}{s^2+4s+5}$

25. $F(s) = \dfrac{s\,e^{-s}}{s^2+4} + \dfrac{e^{-3s}}{(s-2)^4}$

26. $F(s) = \dfrac{e^{-s}}{s^2+4} + \dfrac{(s-1)\,e^{-3s}}{s^4}$

27. $F(s) = \dfrac{(s+1)\,e^{-s}}{s^2+4} + \dfrac{e^{-3s}}{s^4}$

28. Find the Laplace transform of $f(t) = te^t[H(t-1) - H(t-2)]$ by using (a) the definition of the Laplace transform, and (b) a joint application of the first and second shifting theorems.

29. Write the function

$$f(t) = \begin{cases} t, & 0 < t < a, \\ 0, & a < t, \end{cases}$$

in terms of Heaviside's step functions. Then find its transform using (a) the definition of the Laplace transform, and (b) the second shifting theorem.

In Problems 30–33, write the function $f(t)$ in terms of the Heaviside step functions and then find its transform using the second shifting theorem. Check your answer using MATLAB.

30. $f(t) = \begin{cases} t/2, & 0 \le t < 2 \\ 0, & 2 < t \end{cases}$

31. $f(t) = \begin{cases} t, & 0 \le t \le 1 \\ 1, & 1 \le t < 2 \\ 0, & 2 < t \end{cases}$

32. $f(t) = \begin{cases} t, & 0 \le t \le 2 \\ 4 - t, & 2 \le t \le 4 \\ 0, & 4 \le t \end{cases}$

33. $f(t) = \begin{cases} 0, & 0 \le t \le 1 \\ t - 1, & 1 \le t \le 2 \\ 1, & 2 \le t < 3 \\ 0, & 3 < t \end{cases}$

Find $Y(s)$ for the following ordinary differential equations and then use MATLAB to check your work.

34. $y'' + 3y' + 2y = H(t-1); \qquad y(0) = y'(0) = 0$

35. $y'' + 4y = 3H(t-4); \qquad y(0) = 1,\ y'(0) = 0$

36. $y'' + 4y' + 4y = tH(t-2); \qquad y(0) = 0,\ y'(0) = 2$

37. $y'' + 3y' + 2y = e^t H(t-1); \qquad y(0) = y'(0) = 0$

38. $y'' - 3y' + 2y = e^{-t} H(t-2); \qquad y(0) = 2,\ y'(0) = 0$

39. $y'' - 3y' + 2y = t^2 H(t-1); \qquad y(0) = 0,\ y'(0) = 5$

40. $y'' + y = \sin(t)[1 - H(t-\pi)]; \qquad y(0) = y'(0) = 0$

41. $y'' + 3y' + 2y = t + \left[ae^{-(t-a)} - t\right] H(t-a); \qquad y(0) = y'(0) = 0.$

For each of the following functions, find its value at $t = 0$. Then check your answer using the initial-value theorem by hand and using MATLAB.

42. $f(t) = t$

43. $f(t) = \cos(at)$

44. $f(t) = te^{-t}$

45. $f(t) = e^t \sin(3t)$

For each of the following Laplace transforms, state whether you can or cannot apply the final-value theorem. If you can, find the final value by hand and using MATLAB. Check your result by finding the inverse and finding the limit as $t \to \infty$.

46. $F(s) = \dfrac{1}{s-1}$

47. $F(s) = \dfrac{1}{s}$

48. $F(s) = \dfrac{1}{s+1}$

49. $F(s) = \dfrac{s}{s^2+1}$

50. $F(s) = \dfrac{2}{s(s^2+3s+2)}$

51. $F(s) = \dfrac{2}{s(s^2-3s+2)}$

52. Using the fact that

$$e^{-c\xi} = 1 - c\int_0^{\xi} e^{-c\eta}\,d\eta,$$

show[4] that

$$\frac{1}{s}\exp\left(-\frac{asx}{s+b}\right) = \frac{1}{s} - \int_0^{ax} \frac{e^{-\eta}}{s+b}\exp\left(\frac{b\eta}{s+b}\right)\,d\eta$$

if $x > 0$. Therefore, using the fact[5] that

$$\mathcal{L}^{-1}\left(s^{-\nu-1}e^{\alpha/s}\right) = \left(\frac{t}{\alpha}\right)^{\nu/2} I_\nu\left(2\sqrt{\alpha t}\right), \qquad \Re(\nu) > -1,$$

and the first shifting theorem, show

$$\mathcal{L}^{-1}\left[\frac{1}{s}\exp\left(-\frac{asx}{s+b}\right)\right] = 1 - e^{-bt}\int_0^{ax} e^{-\eta} I_0\left(2\sqrt{2t\eta}\right)\,d\eta,$$

where $I_\nu(\cdot)$ is a modified Bessel function of the first kind and order ν introduced in Section 6.5.

12.4 THE LAPLACE TRANSFORM OF A PERIODIC FUNCTION

Periodic functions frequently occur in engineering problems and we shall now show how to calculate their transform. They possess the property that $f(t+T) = f(t)$ for $t > 0$ and equal zero for $t < 0$, where T is the period of the function.

For convenience, let us define a function $x(t)$ that equals zero except over the interval $(0, T)$ where it equals $f(t)$:

$$x(t) = \begin{cases} f(t), & 0 < t < T \\ 0, & T < t. \end{cases} \tag{12.4.1}$$

By definition,

$$F(s) = \int_0^\infty f(t)e^{-st}dt = \int_0^T f(t)e^{-st}dt + \int_T^{2T} f(t)e^{-st}dt + \cdots + \int_{kT}^{(k+1)T} f(t)e^{-st}dt + \cdots. \tag{12.4.2}$$

Now let $z = t - kT$, where $k = 0, 1, 2, \ldots$, in the kth integral and $F(s)$ becomes

$$F(s) = \int_0^T f(z)e^{-sz}\,dz + \int_0^T f(z+T)e^{-s(z+T)}\,dz + \cdots + \int_0^T f(z+kT)e^{-s(z+kT)}\,dz + \cdots. \tag{12.4.3}$$

However,

$$x(z) = f(z) = f(z+T) = \ldots = f(z+kT) = \ldots, \tag{12.4.4}$$

because the range of integration in each integral is from 0 to T. Thus, $F(s)$ becomes

$$F(s) = \int_0^T x(z)e^{-sz}\,dz + e^{-sT}\int_0^T x(z)e^{-sz}\,dz + \cdots + e^{-ksT}\int_0^T x(z)e^{-sz}\,dz + \cdots \tag{12.4.5}$$

---

[4] Liaw, C. H., J. S. P. Wang, R. A. Greenkorn, and K. C. Chao, 1979: Kinetics of fixed-bed absorption: A new solution. *AICHE J.*, **25**, 376–381.

[5] Watson, E. J., 1981: *Laplace Transforms and Applications*. Van Nostrand Reinhold Co., p. 195.

or

$$F(s) = \left(1 + e^{-sT} + e^{-2sT} + \cdots + e^{-ksT} + \cdots\right) X(s). \tag{12.4.6}$$

The first term on the right side of Equation 12.4.6 is a geometric series with common ratio e^{-sT}. If $|e^{-sT}| < 1$, then the series converges and

$$\boxed{F(s) = \frac{X(s)}{1 - e^{-sT}}.} \tag{12.4.7}$$

• **Example 12.4.1**

Let us find the Laplace transform of the square wave with period T:

$$f(t) = \begin{cases} h, & 0 < t < T/2, \\ -h, & T/2 < t < T. \end{cases} \tag{12.4.8}$$

By definition $x(t)$ is

$$x(t) = \begin{cases} h, & 0 < t < T/2, \\ -h, & T/2 < t < T, \\ 0, & T < t. \end{cases} \tag{12.4.9}$$

Then

$$X(s) = \int_0^\infty x(t) e^{-st}\, dt = \int_0^{T/2} h\, e^{-st}\, dt + \int_{T/2}^{T} (-h)\, e^{-st}\, dt \tag{12.4.10}$$

$$= \frac{h}{s}\left(1 - 2e^{-sT/2} + e^{-sT}\right) = \frac{h}{s}\left(1 - e^{-sT/2}\right)^2, \tag{12.4.11}$$

and

$$F(s) = \frac{h\left(1 - e^{-sT/2}\right)^2}{s\left(1 - e^{-sT}\right)} = \frac{h\left(1 - e^{-sT/2}\right)}{s\left(1 + e^{-sT/2}\right)}. \tag{12.4.12}$$

If we multiply numerator and denominator by $\exp(sT/4)$ and recall that $\tanh(u) = (e^u - e^{-u})/(e^u + e^{-u})$, we have that

$$F(s) = \frac{h}{s} \tanh\left(\frac{sT}{4}\right). \tag{12.4.13}$$

□

• **Example 12.4.2**

Let us find the Laplace transform of the periodic function

$$f(t) = \begin{cases} \sin(2\pi t/T), & 0 \le t \le T/2, \\ 0, & T/2 \le t \le T. \end{cases} \tag{12.4.14}$$

By definition $x(t)$ is

$$x(t) = \begin{cases} \sin(2\pi t/T), & 0 \le t \le T/2, \\ 0, & T/2 \le t. \end{cases} \tag{12.4.15}$$

Then

$$X(s) = \int_0^{T/2} \sin\left(\frac{2\pi t}{T}\right) e^{-st}\, dt = \frac{2\pi T}{s^2T^2 + 4\pi^2}\left(1 + e^{-sT/2}\right). \qquad \textbf{(12.4.16)}$$

Hence,

$$F(s) = \frac{X(s)}{1 - e^{-sT}} = \frac{2\pi T}{s^2T^2 + 4\pi^2} \times \frac{1 + e^{-sT/2}}{1 - e^{-sT}} = \frac{2\pi T}{s^2T^2 + 4\pi^2} \times \frac{1}{1 - e^{-sT/2}}. \qquad \textbf{(12.4.17)}$$

Problems

Find the Laplace transform for the following periodic functions:

1. $f(t) = \sin(t), \qquad 0 \le t \le \pi, \qquad f(t) = f(t + \pi)$

2. $f(t) = \begin{cases} \sin(t), & 0 \le t \le \pi, \\ 0, & \pi \le t \le 2\pi, \end{cases} \qquad f(t) = f(t + 2\pi)$

3. $f(t) = \begin{cases} t, & 0 \le t < a, \\ 0, & a < t \le 2a, \end{cases} \qquad f(t) = f(t + 2a)$

4. $f(t) = \begin{cases} 1, & 0 < t < a, \\ 0, & a < t < 2a, \\ -1, & 2a < t < 3a, \\ 0, & 3a < t < 4a, \end{cases} \qquad f(t) = f(t + 4a)$

12.5 INVERSION BY PARTIAL FRACTIONS: HEAVISIDE'S EXPANSION THEOREM

In the previous sections, we devoted our efforts to calculating the Laplace transform of a given function. Obviously, we must have a method for going the other way. Given a transform, we must find the corresponding function. This is often a very formidable task. In the next few sections we shall present some general techniques for the inversion of a Laplace transform.

The first technique involves transforms that we can express as the ratio of two polynomials: $F(s) = q(s)/p(s)$. We shall assume that the order of $q(s)$ is *less* than $p(s)$ and we have divided out any common factor between them. In principle we know that $p(s)$ has n zeros, where n is the order of the $p(s)$ polynomial. Some of the zeros may be complex, some of them may be real, and some of them may be duplicates of other zeros. In the case when $p(s)$ has n simple zeros (nonrepeating roots), a simple method exists for inverting the transform.

We want to rewrite $F(s)$ in the form:

$$F(s) = \frac{a_1}{s - s_1} + \frac{a_2}{s - s_2} + \cdots + \frac{a_n}{s - s_n} = \frac{q(s)}{p(s)}, \qquad \textbf{(12.5.1)}$$

where $s_1, s_2, \ldots, s_n$ are the n simple zeros of $p(s)$. We now multiply both sides of Equation 12.5.1 by $s - s_1$ so that

$$\frac{(s - s_1)q(s)}{p(s)} = a_1 + \frac{(s - s_1)a_2}{s - s_2} + \cdots + \frac{(s - s_1)a_n}{s - s_n}. \qquad \textbf{(12.5.2)}$$

If we set $s = s_1$, the right side of Equation 12.5.2 becomes simply a_1. The left side takes the form 0/0 and there are two cases. If $p(s) = (s - s_1)g(s)$, then $a_1 = q(s_1)/g(s_1)$. If we cannot explicitly factor out $s - s_1$, l'Hôpital's rule gives

$$a_1 = \lim_{s \to s_1} \frac{(s - s_1)q(s)}{p(s)} = \lim_{s \to s_1} \frac{(s - s_1)q'(s) + q(s)}{p'(s)} = \frac{q(s_1)}{p'(s_1)}. \tag{12.5.3}$$

In a similar manner, we can compute all of the coefficients a_k, where $k = 1, 2, \ldots, n$. Therefore,

$$\mathcal{L}^{-1}[F(s)] = \mathcal{L}^{-1}\left[\frac{q(s)}{p(s)}\right] = \mathcal{L}^{-1}\left(\frac{a_1}{s - s_1} + \frac{a_2}{s - s_2} + \cdots + \frac{a_n}{s - s_n}\right) \tag{12.5.4}$$

$$= a_1 e^{s_1 t} + a_2 e^{s_2 t} + \cdots + a_n e^{s_n t}. \tag{12.5.5}$$

This is *Heaviside's expansion theorem*, applicable when $p(s)$ has only simple poles.

• Example 12.5.1

Let us invert the transform $s/[(s+2)(s^2+1)]$. It has three simple poles at $s = -2$ and $s = \pm i$. From our earlier discussion, $q(s) = s$, $p(s) = (s+2)(s^2+1)$, and $p'(s) = 3s^2+4s+1$. Therefore,

$$\mathcal{L}^{-1}\left[\frac{s}{(s+2)(s^2+1)}\right] = \frac{-2}{12 - 8 + 1} e^{-2t} + \frac{i}{-3 + 4i + 1} e^{it} + \frac{-i}{-3 - 4i + 1} e^{-it} \tag{12.5.6}$$

$$= -\frac{2}{5} e^{-2t} + \frac{i}{-2 + 4i} e^{it} - \frac{i}{-2 - 4i} e^{-it} \tag{12.5.7}$$

$$= -\frac{2}{5} e^{-2t} + i \frac{-2 - 4i}{4 + 16} e^{it} - i \frac{-2 + 4i}{4 + 16} e^{-it} \tag{12.5.8}$$

$$= -\frac{2}{5} e^{-2t} + \frac{1}{5} \sin(t) + \frac{2}{5} \cos(t), \tag{12.5.9}$$

where we used $\sin(t) = \frac{1}{2i}(e^{it} - e^{-it})$, and $\cos(t) = \frac{1}{2}(e^{it} + e^{-it})$. □

• Example 12.5.2

Let us invert the transform $1/[(s-1)(s-2)(s-3)]$. There are three simple poles at $s_1 = 1$, $s_2 = 2$, and $s_3 = 3$. In this case, the easiest method for computing a_1, a_2, and a_3 is

$$a_1 = \lim_{s \to 1} \frac{s - 1}{(s - 1)(s - 2)(s - 3)} = \frac{1}{2}, \tag{12.5.10}$$

$$a_2 = \lim_{s \to 2} \frac{s - 2}{(s - 1)(s - 2)(s - 3)} = -1 \tag{12.5.11}$$

and

$$a_3 = \lim_{s \to 3} \frac{s - 3}{(s - 1)(s - 2)(s - 3)} = \frac{1}{2}. \tag{12.5.12}$$

Therefore,

$$\mathcal{L}^{-1}\left[\frac{1}{(s-1)(s-2)(s-3)}\right] = \mathcal{L}^{-1}\left[\frac{a_1}{s-1} + \frac{a_2}{s-2} + \frac{a_3}{s-3}\right] = \tfrac{1}{2}e^t - e^{2t} + \tfrac{1}{2}e^{3t}. \tag{12.5.13}$$

□

Note that for inverting transforms of the form $F(s)e^{-as}$ with $a > 0$, you should use Heaviside's expansion theorem to first invert $F(s)$ and then apply the second shifting theorem.

Let us now find the expansion when we have multiple roots, namely

$$F(s) = \frac{q(s)}{p(s)} = \frac{q(s)}{(s-s_1)^{m_1}(s-s_2)^{m_2}\cdots(s-s_n)^{m_n}}, \qquad \textbf{(12.5.14)}$$

where the order of the denominator, $m_1 + m_2 + \cdots + m_n$, is greater than that for the numerator. Once again we eliminated any common factor between the numerator and denominator. Now we can write $F(s)$ as

$$F(s) = \sum_{k=1}^{n}\sum_{j=1}^{m_k} \frac{a_{kj}}{(s-s_k)^{m_k-j+1}}. \qquad \textbf{(12.5.15)}$$

Multiplying Equation 12.5.15 by $(s-s_k)^{m_k}$,

$$\begin{aligned}\frac{(s-s_k)^{m_k}q(s)}{p(s)} &= a_{k1} + a_{k2}(s-s_k) + \cdots + a_{km_k}(s-s_k)^{m_k-1}\\ &\quad + (s-s_k)^{m_k}\left[\frac{a_{11}}{(s-s_1)^{m_1}} + \cdots + \frac{a_{nm_n}}{s-s_n}\right],\end{aligned} \qquad \textbf{(12.5.16)}$$

where we grouped together into the square-bracketed term all of the terms except for those with a_{kj} coefficients. Taking the limit as $s \to s_k$,

$$a_{k1} = \lim_{s\to s_k} \frac{(s-s_k)^{m_k}q(s)}{p(s)}. \qquad \textbf{(12.5.17)}$$

Let us now take the derivative of Equation 12.5.16,

$$\begin{aligned}\frac{d}{ds}\left[\frac{(s-s_k)^{m_k}q(s)}{p(s)}\right] &= a_{k2} + 2a_{k3}(s-s_k) + \cdots + (m_k-1)a_{km_k}(s-s_k)^{m_k-2}\\ &\quad + \frac{d}{ds}\left\{(s-s_k)^{m_k}\left[\frac{a_{11}}{(s-s_1)^{m_1}} + \cdots + \frac{a_{nm_n}}{s-s_n}\right]\right\}.\end{aligned} \qquad \textbf{(12.5.18)}$$

Taking the limit as $s \to s_k$,

$$a_{k2} = \lim_{s\to s_k} \frac{d}{ds}\left[\frac{(s-s_k)^{m_k}q(s)}{p(s)}\right]. \qquad \textbf{(12.5.19)}$$

In general,

$$a_{kj} = \lim_{s\to s_k} \frac{1}{(j-1)!}\frac{d^{j-1}}{ds^{j-1}}\left[\frac{(s-s_k)^{m_k}q(s)}{p(s)}\right], \qquad \textbf{(12.5.20)}$$

and by direct inversion,

$$f(t) = \sum_{k=1}^{n}\sum_{j=1}^{m_k} \frac{a_{kj}}{(m_k-j)!}t^{m_k-j}e^{s_k t}. \qquad \textbf{(12.5.21)}$$

• **Example 12.5.3**

Let us find the inverse of

$$F(s) = \frac{s}{(s+2)^2(s^2+1)}. \tag{12.5.22}$$

We first note that the denominator has simple zeros at $s = \pm i$ and a repeated root at $s = -2$. Therefore,

$$F(s) = \frac{A}{s-i} + \frac{B}{s+i} + \frac{C}{s+2} + \frac{D}{(s+2)^2}, \tag{12.5.23}$$

where

$$A = \lim_{s\to i}\,(s-i)F(s) = \tfrac{1}{6+8i}, \tag{12.5.24}$$

$$B = \lim_{s\to -i}\,(s+i)F(s) = \tfrac{1}{6-8i}, \tag{12.5.25}$$

$$C = \lim_{s\to -2}\,\frac{d}{ds}\Big[(s+2)^2F(s)\Big] = \lim_{s\to -2}\,\frac{d}{ds}\left(\frac{s}{s^2+1}\right) = -\tfrac{3}{25}, \tag{12.5.26}$$

and

$$D = \lim_{s\to -2}\,(s+2)^2F(s) = -\tfrac{2}{5}. \tag{12.5.27}$$

Thus,

$$f(t) = \tfrac{1}{6+8i}e^{it} + \tfrac{1}{6-8i}e^{-it} - \tfrac{3}{25}e^{-2t} - \tfrac{2}{5}te^{-2t} = \tfrac{3}{25}\cos(t) + \tfrac{4}{25}\sin(t) - \tfrac{3}{25}e^{-2t} - \tfrac{10}{25}te^{-2t}. \tag{12.5.28}$$

□

In Section 12.9 we shall see that we can invert transforms just as easily with the residue theorem. Let us now find the inverse of

$$F(s) = \frac{cs + (ca - \omega d)}{(s+a)^2 + \omega^2} = \frac{cs + (ca - \omega d)}{(s+a-\omega i)(s+a+\omega i)} \tag{12.5.29}$$

by Heaviside's expansion theorem. Then

$$F(s) = \frac{c+di}{2(s+a-\omega i)} + \frac{c-di}{2(s+a+\omega i)} = \frac{\sqrt{c^2+d^2}e^{\theta i}}{2(s+a-\omega i)} + \frac{\sqrt{c^2+d^2}e^{-\theta i}}{2(s+a+\omega i)}, \tag{12.5.30}$$

where $\theta = \tan^{-1}(d/c)$. Note that we must choose θ so that it gives the correct sign for c and d.

Taking the inverse of Equation 12.5.30,

$$f(t) = \tfrac{1}{2}\sqrt{c^2+d^2}e^{-at+\omega ti+\theta i} + \tfrac{1}{2}\sqrt{c^2+d^2}e^{-at-\omega ti-\theta i} = \sqrt{c^2+d^2}e^{-at}\cos(\omega t+\theta). \tag{12.5.31}$$

Equation 12.5.31 is the amplitude/phase form of the inverse of Equation 12.5.29. It is particularly popular with electrical engineers.

• **Example 12.5.4**

Let us express the inverse of

$$F(s) = \frac{8s-3}{s^2+4s+13} \tag{12.5.32}$$

in the amplitude/phase form.

Starting with

$$F(s) = \frac{8s-3}{(s+2-3i)(s+2+3i)} = \frac{4+19i/6}{s+2-3i} + \frac{4-19i/6}{s+2+3i} \tag{12.5.33}$$

$$= \frac{5.1017e^{38.3675^\circ i}}{s+2-3i} + \frac{5.1017e^{-38.3675^\circ i}}{s+2+3i}, \tag{12.5.34}$$

or

$$f(t) = 5.1017e^{-2t+3it+38.3675^\circ i} + 5.1017e^{-2t-3it-38.3675^\circ i} \tag{12.5.35}$$

$$= 10.2034e^{-2t}\cos(3t+38.3675^\circ). \tag{12.5.36}$$

□

• **Example 12.5.5: The design of film projectors**

For our final example we anticipate future work. The primary use of Laplace transforms is the solution of differential equations. In this example we illustrate this technique that includes Heaviside's expansion theorem in the form of amplitude and phase.

This problem[6] arose in the design of projectors for motion pictures. An early problem was ensuring that the speed at which the film passed the electric eye remained essentially constant; otherwise, a frequency modulation of the reproduced sound resulted. Figure 12.5.1(A) shows a diagram of the projector. Many will remember this design from their days as a school projectionist. In this section we shall show that this particular design filters out variations in the film speed caused by irregularities either in the driving-gear trains or in the engagement of the sprocket teeth with the holes in the film.

Let us now focus on the film head - a hollow drum of small moment of inertia J_1. See Figure 12.5.1(B). Within it there is a concentric inner flywheel of moment of inertia J_2, where $J_2 \gg J_1$. The remainder of the space within the drum is filled with oil. The inner flywheel rotates on precision ball bearings on the drum shaft. The only coupling between the drum and flywheel is through fluid friction and the very small friction in the ball bearings. The flection of the film-loops between the drum head and idler pulleys provides the spring restoring force for the system as the film runs rapidly through the system.

From Figure 12.5.1 the dynamical equations governing the outer case and inner flywheel are (1) the rate of change of the outer casing of the film head equals the frictional torque given to the casing from the inner flywheel plus the restoring torque due to the flection of the film, and (2) the rate of change of the inner flywheel equals the negative of the frictional torque given to the outer casing by the inner flywheel.

Assuming that the frictional torque between the two flywheels is proportional to the difference in their angular velocities, the frictional torque given to the casing from the inner flywheel is $B(\omega_2-\omega_1)$, where B is the frictional resistance, ω_1 and ω_2 are the deviations of the drum and inner flywheel from their normal angular velocities, respectively. If r is the ratio of the diameter of the winding sprocket to the diameter of the drum, the restoring torque due to the flection of the film and its corresponding angular twist equals $K\int_0^t (r\omega_0-\omega_1)\,d\tau$, where K is the rotational stiffness and ω_0 is the deviation of the winding sprocket from its

[6] Cook, E. D., 1935: The technical aspects of the high-fidelity reproducer. *J. Soc. Motion Pict. Eng.*, **25**, 289–312.

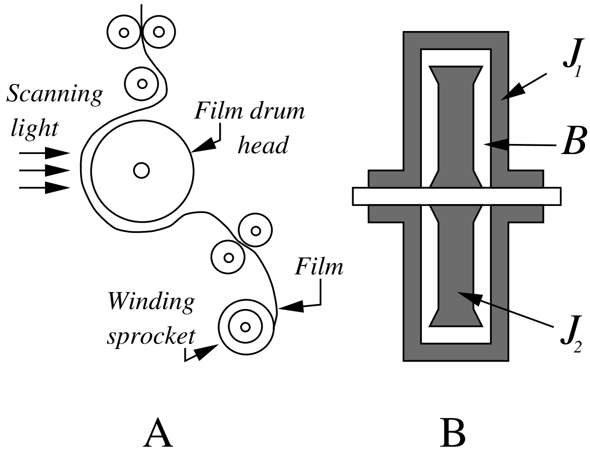


Figure 12.5.1: (A) The schematic for the scanning light in a motion-picture projector and (B) interior of the film drum head.

normal angular velocity. The quantity $r\omega_0$ gives the angular velocity at which the film is running through the projector because the winding sprocket is the mechanism that pulls the film. Consequently, the equations governing this mechanical system are

$$J_1 \frac{d\omega_1}{dt} = K \int_0^t (r\omega_0 - \omega_1)\, d\tau + B(\omega_2 - \omega_1), \qquad (\mathbf{12.5.37})$$

and

$$J_2 \frac{d\omega_2}{dt} = -B(\omega_2 - \omega_1). \qquad (\mathbf{12.5.38})$$

With the winding sprocket, the drum, and the flywheel running at their normal uniform angular velocities, let us assume that the winding sprocket introduces a disturbance equivalent to a unit increase in its angular velocity for 0.15 second, followed by the resumption of its normal velocity. It is assumed that the film in contact with the drum cannot slip. The initial conditions are $\omega_1(0) = \omega_2(0) = 0$.

Taking the Laplace transform of Equation 12.5.37 and Equation 12.5.38 and using Equation 12.1.18,

$$\left(J_1 s + B + \frac{K}{s}\right)\Omega_1(s) - B\Omega_2(s) = \frac{rK}{s}\Omega_0(s) = rK\mathcal{L}\left[\int_0^t \omega_0(\tau)\, d\tau\right], \qquad (\mathbf{12.5.39})$$

and

$$-B\Omega_1(s) + (J_2 s + B)\Omega_2(s) = 0. \qquad (\mathbf{12.5.40})$$

The solution of Equation 12.5.39 and Equation 12.5.40 for $\Omega_1(s)$ is

$$\Omega_1(s) = \frac{rK}{J_1} \quad \frac{(s + a_0)\Omega_0(s)}{s^3 + b_2 s^2 + b_1 s + b_0}, \qquad (\mathbf{12.5.41})$$

where typical values[7] are

$$\frac{rK}{J_1} = 90.8, \quad a_0 = \frac{B}{J_2} = 1.47, \quad b_0 = \frac{BK}{J_1 J_2} = 231, \qquad (\mathbf{12.5.42})$$

[7] $J_1 = 1.84 \times 10^4$ dyne cm sec^2 per radian, $J_2 = 8.43 \times 10^4$ dyne cm sec^2 per radian, $B = 12.4 \times 10^4$ dyne cm sec per radian, $K = 2.89 \times 10^6$ dyne cm per radian, and $r = 0.578$.

$$b_1 = \frac{K}{J_1} = 157, \quad \text{and} \quad b_2 = \frac{B(J_1 + J_2)}{J_1 J_2} = 8.20. \tag{12.5.43}$$

The transform $\Omega_1(s)$ has three simple poles located at $s_1 = -1.58, s_2 = -3.32 + 11.6i$, and $s_3 = -3.32 - 11.6i$.

Because the sprocket angular velocity deviation $\omega_0(t)$ is a pulse of unit amplitude and 0.15 second duration, we express it as the difference of two Heaviside step functions

$$\omega_0(t) = H(t) - H(t - 0.15). \tag{12.5.44}$$

Its Laplace transform is

$$\Omega_0(s) = \frac{1}{s} - \frac{1}{s} e^{-0.15s} \tag{12.5.45}$$

so that Equation 12.5.41 becomes

$$\Omega_1(s) = \frac{rK}{J_1} \frac{(s + a_0)}{s(s - s_1)(s - s_2)(s - s_3)} \left(1 - e^{-0.15s}\right). \tag{12.5.46}$$

The inversion of Equation 12.5.46 follows directly from the second shifting theorem and Heaviside's expansion theorem, or

$$\begin{aligned} \omega_1(t) = {} & K_0 + K_1 e^{s_1 t} + K_2 e^{s_2 t} + K_3 e^{s_3 t} \\ & - [K_0 + K_1 e^{s_1 (t-0.15)} + K_2 e^{s_2 (t-0.15)} + K_3 e^{s_3 (t-0.15)}] H(t - 0.15), \end{aligned} \tag{12.5.47}$$

where

$$K_0 = \frac{rK}{J_1} \left. \frac{s + a_0}{(s - s_1)(s - s_2)(s - s_3)} \right|_{s=0} = 0.578, \tag{12.5.48}$$

$$K_1 = \frac{rK}{J_1} \left. \frac{s + a_0}{s(s - s_2)(s - s_3)} \right|_{s=s_1} = 0.046, \tag{12.5.49}$$

$$K_2 = \frac{rK}{J_1} \left. \frac{s + a_0}{s(s - s_1)(s - s_3)} \right|_{s=s_2} = 0.326 e^{165^\circ i}, \tag{12.5.50}$$

and

$$K_3 = \frac{rK}{J_1} \left. \frac{s + a_0}{s(s - s_1)(s - s_2)} \right|_{s=s_3} = 0.326 e^{-165^\circ i}. \tag{12.5.51}$$

Using Euler's identity $\cos(t) = (e^{it} + e^{-it})/2$, we can write Equation 12.5.47 as

$$\begin{aligned} \omega_1(t) = {} & 0.578 + 0.046 e^{-1.58t} + 0.652 e^{-3.32t} \cos(11.6t + 165^\circ) \\ & - \{0.578 + 0.046 e^{-1.58(t-0.15)} + 0.652 e^{-3.32(t-0.15)} \\ & \times \cos[11.6(t - 0.15) + 165^\circ]\} H(t - 0.15). \end{aligned} \tag{12.5.52}$$

Equation 12.5.52 is plotted in Figure 12.5.2. Note that fluctuations in $\omega_1(t)$ are damped out by the particular design of this film projector. Because this mechanical device dampens unwanted fluctuations (or noise) in the motion-picture projector, this particular device is an example of a *mechanical filter*.

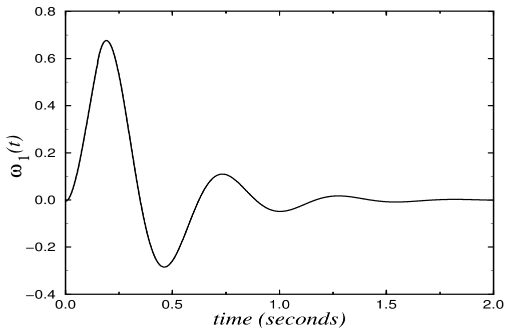


Figure 12.5.2: The deviation $\omega_1(t)$ of a film drum head from its uniform angular velocity when the sprocket angular velocity is perturbed by a unit amount for the duration of 0.15 second.

Problems

Use Heaviside's expansion theorem to find the inverse of the following Laplace transforms:

1. $F(s) = \dfrac{1}{s^2 + 3s + 2}$

2. $F(s) = \dfrac{s+3}{(s+4)(s-2)}$

3. $F(s) = \dfrac{s-4}{(s+2)(s+1)(s-3)}$

4. $F(s) = \dfrac{s-3}{(s^2+4)(s+1)}$

Find the inverse of the following transforms and express them in amplitude/phase form:

5. $F(s) = \dfrac{1}{s^2 + 4s + 5}$

6. $F(s) = \dfrac{1}{s^2 + 6s + 13}$

7. $F(s) = \dfrac{2s-5}{s^2 + 16}$

8. $F(s) = \dfrac{1}{s(s^2 + 2s + 2)}$

9. $F(s) = \dfrac{s+2}{s(s^2+4)}$

12.6 CONVOLUTION

In this section we turn to a fundamental concept in Laplace transforms: convolution. We shall restrict ourselves to its use in finding the inverse of a transform when that transform consists of the *product* of two simpler transforms. In subsequent sections we will use it to solve ordinary differential equations.

We begin by formally introducing the mathematical operation of the *convolution product*

$$f(t) * g(t) = \int_0^t f(t-x)g(x)\,dx = \int_0^t f(x)g(t-x)\,dx. \qquad \textbf{(12.6.1)}$$

In most cases the operations required by Equation 12.6.1 are straightforward.

• Example 12.6.1

Let us find the convolution between $\cos(t)$ and $\sin(t)$.

$$\cos(t) * \sin(t) = \int_0^t \sin(t-x)\cos(x)\,dx = \tfrac{1}{2}\int_0^t [\sin(t) + \sin(t-2x)]\,dx \qquad \textbf{(12.6.2)}$$

$$= \tfrac{1}{2}\int_0^t \sin(t)\,dx + \tfrac{1}{2}\int_0^t \sin(t-2x)\,dx \qquad (\mathbf{12.6.3})$$

$$= \tfrac{1}{2}\sin(t)\,x\Big|_0^t + \tfrac{1}{4}\cos(t-2x)\Big|_0^t = \tfrac{1}{2}t\sin(t). \qquad (\mathbf{12.6.4})$$

□

• Example 12.6.2

Similarly, the convolution between t^2 and $\sin(t)$ is

$$t^2 * \sin(t) = \int_0^t (t-x)^2 \sin(x)\,dx \qquad (\mathbf{12.6.5})$$

$$= -(t-x)^2\cos(x)\Big|_0^t - 2\int_0^t (t-x)\cos(x)\,dx \qquad (\mathbf{12.6.6})$$

$$= t^2 - 2(t-x)\sin(x)\Big|_0^t - 2\int_0^t \sin(x)\,dx \qquad (\mathbf{12.6.7})$$

$$= t^2 + 2\cos(t) - 2 \qquad (\mathbf{12.6.8})$$

by integration by parts. □

• Example 12.6.3

Consider now the convolution between e^t and the discontinuous function $H(t-1) - H(t-2)$:

$$e^t * [H(t-1) - H(t-2)] = \int_0^t e^{t-x}[H(x-1) - H(x-2)]\,dx \qquad (\mathbf{12.6.9})$$

$$= e^t \int_0^t e^{-x}[H(x-1) - H(x-2)]\,dx. \qquad (\mathbf{12.6.10})$$

In order to evaluate the integral, Equation 12.6.10, we must examine various cases. If $t < 1$, then both of the step functions equal zero and the convolution equals zero. However, when $1 < t < 2$, the first step function equals one while the second equals zero as the dummy variable x runs between 1 and t. Therefore,

$$e^t * [H(t-1) - H(t-2)] = e^t \int_1^t e^{-x} dx = e^{t-1} - 1, \qquad (\mathbf{12.6.11})$$

because the portion of the integral from zero to one equals zero. Finally, when $t > 2$, the integrand is only nonzero for that portion of the integration when $1 < x < 2$. Consequently,

$$e^t * [H(t-1) - H(t-2)] = e^t \int_1^2 e^{-x} dx = e^{t-1} - e^{t-2}. \qquad (\mathbf{12.6.12})$$

Thus, the convolution of e^t with the pulse $H(t-1) - H(t-2)$ is

$$e^t * [H(t-1) - H(t-2)] = \begin{cases} 0, & 0 \le t \le 1, \\ e^{t-1} - 1, & 1 \le t \le 2, \\ e^{t-1} - e^{t-2}, & 2 \le t. \end{cases} \qquad (\mathbf{12.6.13})$$

MATLAB can also be used to find the convolution of two functions. For example, in the present case the commands

```
syms x t positive
int('exp(t-x)*(Heaviside(x-1)-Heaviside(x-2))',x,0,t)
```

yield

```
ans =
-Heaviside(t-1)+Heaviside(t-1)*exp(t-1)+Heaviside(t-2)
    -Heaviside(t-2)*exp(t-2)
```

□

The reason why we introduced convolution stems from the following fundamental theorem (often called *Borel's theorem*[8]). If

$$w(t) = u(t) * v(t) \tag{12.6.14}$$

then

$$W(s) = U(s)V(s). \tag{12.6.15}$$

In other words, we can invert a complicated transform by convoluting the inverses to two simpler functions. The proof is as follows:

$$W(s) = \int_0^\infty \left[\int_0^t u(x)v(t-x)\,dx\right] e^{-st}dt \tag{12.6.16}$$

$$= \int_0^\infty \left[\int_x^\infty u(x)v(t-x)e^{-st}dt\right] dx \tag{12.6.17}$$

$$= \int_0^\infty u(x) \left[\int_0^\infty v(r)e^{-s(r+x)}dr\right] dx \tag{12.6.18}$$

$$= \left[\int_0^\infty u(x)e^{-sx}dx\right] \left[\int_0^\infty v(r)e^{-sr}dr\right] = U(s)V(s), \tag{12.6.19}$$

where $t = r + x$. □

• Example 12.6.4

Let us find the inverse of the transform

$$\frac{s}{(s^2+1)^2} = \frac{s}{s^2+1} \times \frac{1}{s^2+1} = \mathcal{L}[\cos(t)]\mathcal{L}[\sin(t)] = \mathcal{L}[\cos(t) * \sin(t)] = \mathcal{L}[\tfrac{1}{2}t\sin(t)] \tag{12.6.20}$$

from Example 12.6.1. □

• Example 12.6.5

Let us find the inverse of the transform

$$\frac{1}{(s^2+a^2)^2} = \frac{1}{a^2}\left(\frac{a}{s^2+a^2} \times \frac{a}{s^2+a^2}\right) = \frac{1}{a^2}\mathcal{L}[\sin(at)]\mathcal{L}[\sin(at)]. \tag{12.6.21}$$

[8] Borel, É., 1901: *Leçons sur les séries divergentes*. Gauthier-Villars, p. 104.

Therefore,

$$\mathcal{L}^{-1}\left[\frac{1}{(s^2+a^2)^2}\right] = \frac{1}{a^2}\int_0^t \sin[a(t-x)]\sin(ax)\,dx \tag{12.6.22}$$

$$= \frac{1}{2a^2}\int_0^t \cos[a(t-2x)]\,dx - \frac{1}{2a^2}\int_0^t \cos(at)\,dx \tag{12.6.23}$$

$$= -\frac{1}{4a^3}\sin[a(t-2x)]\Big|_0^t - \frac{1}{2a^2}\cos(at)\,x\Big|_0^t \tag{12.6.24}$$

$$= \frac{1}{2a^3}[\sin(at) - at\cos(at)]. \tag{12.6.25}$$

□

• Example 12.6.6

Let us use the results from Example 12.6.3 to verify the convolution theorem.

We begin by rewriting Equation 12.6.13 in terms of the Heaviside step functions. Using the method outline in Example 12.2.1,

$$f(t)*g(t) = \left(e^{t-1}-1\right)H(t-1) + \left(1-e^{t-2}\right)H(t-2). \tag{12.6.26}$$

Employing the second shifting theorem,

$$\mathcal{L}[f*g] = \frac{e^{-s}}{s-1} - \frac{e^{-s}}{s} + \frac{e^{-2s}}{s} - \frac{e^{-2s}}{s-1} \tag{12.6.27}$$

$$= \frac{e^{-s}}{s(s-1)} - \frac{e^{-2s}}{s(s-1)} = \frac{1}{s-1}\left(\frac{e^{-s}}{s} - \frac{e^{-2s}}{s}\right) \tag{12.6.28}$$

$$= \mathcal{L}[e^t]\mathcal{L}[H(t-1) - H(t-2)] \tag{12.6.29}$$

and the convolution theorem holds true. If we had not rewritten Equation 12.6.13 in terms of step functions, we could still have found $\mathcal{L}[f*g]$ from the definition of the Laplace transform.

Problems

Verify the following convolutions and then show that the convolution theorem is true. Use MATLAB to check your answer.

1. $1*1=t$
2. $1*\cos(at)=\sin(at)/a$
3. $1*e^t=e^t-1$
4. $t*t=t^3/6$
5. $t*\sin(t)=t-\sin(t)$
6. $t*e^t=e^t-t-1$
7. $t^2*\sin(at)=\dfrac{t^2}{a}-\dfrac{4}{a^3}\sin^2\left(\dfrac{at}{2}\right)$
8. $t*H(t-1)=\frac{1}{2}(t-1)^2H(t-1)$
9. $H(t-a)*H(t-b)=(t-a-b)H(t-a-b)$

10. $t * [H(t) - H(t-2)] = \frac{t^2}{2} - \frac{(t-2)^2}{2} H(t-2)$

Use the convolution theorem to invert the following functions:

11. $F(s) = \frac{1}{s^2(s-1)}$

12. $F(s) = \frac{1}{s^2(s+a)^2}$

13. Given[9]

$$\mathcal{L}^{-1}\left[\frac{e^{\alpha/s}}{s}\right] = I_0\left(2\sqrt{\alpha t}\right), \qquad \alpha > 0,$$

show that

$$\frac{e^{a/s}}{s-1} = \left(1 + \frac{1}{s-1}\right)\frac{e^{a/s}}{s}, \qquad a > 0,$$

and

$$\mathcal{L}^{-1}\left(\frac{e^{a/s}}{s-1}\right) = \left[\delta(t) + e^t\right] * I_0\left(2\sqrt{at}\right) = e^t + e^t \int_0^{\sqrt{4at}} \exp\left(-\frac{\tau^2}{4a}\right) I_1(\tau)\, d\tau,$$

where $I_n(\cdot)$ denotes a modified Bessel function of the first kind and order n, which was introduced in Section 6.5. There we will show that $I_0'(\tau) = I_1(\tau)$.

14. Prove that the convolution of two Dirac delta functions is a Dirac delta function.

12.7 INTEGRAL EQUATIONS

An *integral equation* contains the dependent variable under an integral sign. The convolution theorem provides an excellent tool for solving a very special class of these equations, *Volterra equation of the second kind*:[10]

$$f(t) - \int_0^t K[t, x, f(x)]\, dx = g(t), \qquad 0 \le t \le T. \tag{12.7.1}$$

These equations appear in history-dependent problems, such as epidemics,[11] vibration problems,[12] and viscoelasticity.[13]

---

[9] Watson, op. cit., p. 195.

[10] Fock, V., 1924: Über eine Klasse von Integralgleichungen. *Math. Z.*, **21**, 161–173; Koizumi, S., 1931: On Heaviside's operational solution of a Volterra's integral equation when its nucleus is a function of $(x-\xi)$. *Philos. Mag., Ser. 7*, **11**, 432–441.

[11] Wang, F. J. S., 1978: Asymptotic behavior of some deterministic epidemic models. *SIAM J. Math. Anal.*, **9**, 529–534.

[12] Lin, S. P., 1975: Damped vibration of a string. *J. Fluid Mech.*, **72**, 787–797.

[13] Rogers, T. G., and E. H. Lee, 1964: The cylinder problem in viscoelastic stress analysis. *Q. Appl. Math.*, **22**, 117–131.

• **Example 12.7.1**

Let us find $f(t)$ from the integral equation

$$f(t) = 4t - 3\int_0^t f(x)\sin(t-x)\,dx. \tag{12.7.2}$$

The integral in Equation 12.7.2 is such that we can use the convolution theorem to find its Laplace transform. Then, because $\mathcal{L}[\sin(t)] = 1/(s^2+1)$, the convolution theorem yields

$$\mathcal{L}\left[\int_0^t f(x)\sin(t-x)\,dx\right] = \frac{F(s)}{s^2+1}. \tag{12.7.3}$$

Therefore, the Laplace transform converts Equation 12.7.2 into

$$F(s) = \frac{4}{s^2} - \frac{3F(s)}{s^2+1}. \tag{12.7.4}$$

Solving for $F(s)$,

$$F(s) = \frac{4(s^2+1)}{s^2(s^2+4)}. \tag{12.7.5}$$

By partial fractions, or by inspection,

$$F(s) = \frac{1}{s^2} + \frac{3}{s^2+4}. \tag{12.7.6}$$

Therefore, inverting term by term,

$$f(t) = t + \tfrac{3}{2}\sin(2t). \tag{12.7.7}$$

Note that the integral equation

$$f(t) = 4t - 3\int_0^t f(t-x)\sin(x)\,dx \tag{12.7.8}$$

also has the same solution. □

• **Example 12.7.2**

Let us solve the equation

$$f'(t) + \alpha^2\int_0^t f(\tau)\,d\tau = B - C\cos(\omega t), \qquad f(0) = 0. \tag{12.7.9}$$

Again the integral is one of the convolution type; it differs from the previous example in that it includes a derivative. Taking the Laplace transform of Equation 12.7.9,

$$sF(s) - f(0) + \frac{\alpha^2 F(s)}{s} = \frac{B}{s} - \frac{sC}{s^2+\omega^2}. \tag{12.7.10}$$

Because $f(0) = 0$, Equation 12.7.10 simplifies to

$$(s^2+\alpha^2)F(s) = B - \frac{Cs^2}{s^2+\omega^2}. \tag{12.7.11}$$

Solving for $F(s)$,

$$F(s) = \frac{B}{s^2+\alpha^2} - \frac{Cs^2}{(s^2+\alpha^2)(s^2+\omega^2)}. \tag{12.7.12}$$

Using partial fractions to invert Equation 12.7.12,

$$f(t) = \left(\frac{B}{\alpha} + \frac{\alpha C}{\omega^2-\alpha^2}\right)\sin(\alpha t) - \frac{\omega C}{\omega^2-\alpha^2}\sin(\omega t). \tag{12.7.13}$$

□

• Example 12.7.3

Let us solve[14] the integral equation

$$f(t) = \frac{a}{2(1+2a)}\int_0^t f(t-x)f(x)\,dx + e^{-t}. \tag{12.7.14}$$

Taking the Laplace transform of Equation 12.7.14, we obtain

$$F(s) = \frac{a\,F^2(s)}{2(1+2a)} + \frac{1}{s+1}. \tag{12.7.15}$$

Solving for $F(s)$ so that $F(s)\to 0$ as $s\to\infty$, we have

$$F(s) = \frac{2a+1}{a} - \frac{2a+1}{a}\sqrt{\frac{(2a+1)(s+1)-2a}{(2a+1)(s+1)}} = \frac{2a+1}{a} - \frac{\sqrt{2a+1}}{a}\sqrt{\frac{(2a+1)s+1}{s+1}}. \tag{12.7.16}$$

Taking the inverse of Equation 12.7.16,

$$f(t) = \frac{2a+1}{a}\,\delta(t) - \frac{\sqrt{2a+1}}{a}g(t), \tag{12.7.17}$$

where $g(t)$ is the inverse of the Laplace transform $G(s)$,

$$G(s) = \sqrt{\frac{(2a+1)s+1}{s+1}} = \sqrt{2a+1}\,\frac{s+1/(1+2a)}{\sqrt{s+1}\sqrt{s+1/(2a+1)}} \tag{12.7.18}$$

$$= \sqrt{2a+1}\,sH(s) + \frac{H(s)}{\sqrt{2a+1}} \tag{12.7.19}$$

and

$$H(s) = \frac{1}{\sqrt{s+1}\sqrt{s+1/(2a+1)}}. \tag{12.7.20}$$

Taking the inverse of $H(s)$, we find that

$$h(t) = \exp\left(-\frac{a+1}{2a+1}t\right) I_0\left(\frac{at}{2a+1}\right) \tag{12.7.21}$$

[14] Hounslow, M. J., 1990: A discretized population balance for continuous systems at steady state. *AICHE J.*, **36**, 106–116.

and

$$h'(t) = -\frac{a+1}{2a+1}\exp\left(-\frac{a+1}{2a+1}t\right) I_0\left(\frac{at}{2a+1}\right) + \frac{a}{2a+1}\exp\left(-\frac{a+1}{2a+1}t\right) I_1\left(\frac{at}{2a+1}\right), \tag{12.7.22}$$

where $I_0(\cdot)$ and $I_1(\cdot)$ are modified Bessel functions of the first kind. See Section 6.5.

Because $sH(s) = \mathcal{L}[h'(t)] + h(0)$ and $h(0) = 1$, $h'(t) = \mathcal{L}^{-1}[sH(s)] - \delta(t)$ or $\mathcal{L}^{-1}[sH(s)] = h'(t) + \delta(t)$. Then,

$$g(t) = \sqrt{2a+1}\left[h'(t) + \frac{h(t)}{2a+1} + \delta(t)\right] \tag{12.7.23}$$

$$= \sqrt{2a+1}\left[\delta(t) + \frac{a}{2a+1}\exp\left(-\frac{a+1}{2a+1}t\right) I_1\left(\frac{at}{2a+1}\right) - \frac{a}{2a+1}\exp\left(-\frac{a+1}{2a+1}t\right) I_0\left(\frac{at}{2a+1}\right)\right]. \tag{12.7.24}$$

Finally, substituting Equation 12.7.24 into Equation 12.7.17,

$$f(t) = \exp\left(-\frac{a+1}{2a+1}t\right)\left[I_0\left(\frac{at}{2a+1}\right) - I_1\left(\frac{at}{2a+1}\right)\right]. \tag{12.7.25}$$

Problems

Solve the following integral equations:

1. $f(t) = 1 + 2\int_0^t f(t-x)e^{-2x}\,dx$

2. $f(t) = 1 + \int_0^t f(x)\sin(t-x)\,dx$

3. $f(t) = t + \int_0^t f(t-x)e^{-x}\,dx$

4. $f(t) = 4t^2 - \int_0^t f(t-x)e^{-x}\,dx$

5. $f(t) = t^3 + \int_0^t f(x)\sin(t-x)\,dx$

6. $f(t) = 8t^2 - 3\int_0^t f(x)\sin(t-x)\,dx$

7. $f(t) = t^2 - 2\int_0^t f(t-x)\sinh(2x)\,dx$

8. $f(t) = 1 + 2\int_0^t f(t-x)\cos(x)\,dx$

9. $f(t) = e^{2t} + 2\int_0^t f(t-x)\cos(x)\,dx$

10. $f(t) = t^2 + \int_0^t f(x)\sin(t-x)\,dx$

11. $f(t) = e^{-t} - 2\int_0^t f(x)\cos(t-x)\,dx$

12. $f(t) = 6t + 4\int_0^t f(x)(x-t)^2\,dx$

13. $f(t) = a\sqrt{t} - \int_0^t \frac{f(t-x)}{\sqrt{x}}\,dx$

14. Solve the following equation for $f(t)$ with the condition that $f(0) = 4$:

$$f'(t) = t + \int_0^t f(t-x)\cos(x)\,dx.$$

15. Solve the following equation for $f(t)$ with the condition that $f(0) = 0$:

$$f'(t) = \sin(t) + \int_0^t f(t-x)\cos(x)\,dx.$$

16. During a study of nucleation involving idealized active sites along a boiling surface, Marto and Rohsenow[15] solved the integral equation

$$A = B\sqrt{t} + C \int_0^t \frac{x'(\tau)}{\sqrt{t-\tau}}\, d\tau$$

to find the position $x(t)$ of the liquid/vapor interface. If A, B, and C are constants and $x(0) = 0$, find the solution for them.

17. Solve the following equation for $x(t)$ with the condition that $x(0) = 0$:

$$x(t) + t = \frac{1}{c\sqrt{\pi}} \int_0^t \frac{x'(\tau)}{\sqrt{t-\tau}}\, d\tau,$$

where c is constant.

18. During a study of the temperature $f(t)$ of a heat reservoir attached to a semi-infinite heat-conducting rod, Huber[16] solved the integral equation

$$f'(t) = \alpha - \frac{\beta}{\sqrt{\pi}} \int_0^t \frac{f'(\tau)}{\sqrt{t-\tau}}\, d\tau,$$

where α and β are constants and $f(0) = 0$. Find $f(t)$ for him. Hint:

$$\frac{\alpha}{s^{3/2}(s^{1/2}+\beta)} = \frac{\alpha}{s(s-\beta^2)} - \frac{\alpha\beta}{s^{3/2}(s-\beta^2)}.$$

19. During the solution of a diffusion problem, Zhdanov, Chikhachev, and Yavlinskii[17] solved an integral equation similar to

$$\int_0^t f(\tau)\left[1 - \text{erf}\left(a\sqrt{t-\tau}\,\right)\right] d\tau = at,$$

where $\text{erf}(x) = \dfrac{2}{\sqrt{\pi}} \displaystyle\int_0^x e^{-y^2}\, dy$ is the error function. What should they have found? Hint: You will need to prove that

$$\mathcal{L}\left[t\,\text{erf}(a\sqrt{t}\,) - \frac{1}{2a^2}\text{erf}(a\sqrt{t}\,) + \frac{\sqrt{t}}{a\sqrt{\pi}} e^{-a^2 t}\right] = \frac{a}{s^2\sqrt{s+a^2}}.$$

20. The *Laguerre polynomial*[18]

$$y(t) = L_n(t) = \frac{e^t}{n!}\frac{d^n}{dt^n}\left(t^n e^{-t}\right), \qquad n = 0, 1, 2, 3, \ldots$$

---

[15] Marto, P. J., and W. M. Rohsenow, 1966: Nucleate boiling instability of alkali metals. *J. Heat Transfer*, **88**, 183–193.

[16] Huber, A., 1934: Eine Methode zur Bestimmung der Wärme- und Temperaturleitfähigkeit. *Monatsh. Math. Phys.*, **41**, 35–42.

[17] Zhdanov, S. K., A. S. Chikhachev, and Yu. N. Yavlinskii, 1976: Diffusion boundary-value problem for regions with moving boundaries and conservation of particles. *Sov. Phys. Tech. Phys.*, **21**, 883–884.

[18] See Section 5.3 in Andrews, L. C., 1985: *Special Functions for Engineers and Applied Mathematicians*. MacMillan, 357 pp.

satisfies the ordinary differential equation

$$ty'' + (1-t)y' + ny = (ty')' - ty' + ny = 0,$$

with $y(0) = 1$ and $y'(0) = -n$.

Step 1: Using Equation 12.1.18 and Equation 12.3.20, show that the Laplace transformed version of this differential equation is

$$Y'(s) = \frac{n+1-s}{s(s-1)} Y(s) = \frac{n}{s-1} Y(s) - \frac{n+1}{s} Y(s),$$

where $Y(s)$ is the Laplace transform of $y(t)$.

Step 2: Using Equation 12.3.20 and the convolution theorem, show that Laguerre polynomials are the solution to the integral equation

$$ty(t) = (n+1) \int_0^t y(\tau)\, d\tau - ne^t \int_0^t y(\tau)\, e^{-\tau}\, d\tau.$$

12.8 SOLUTION OF LINEAR DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

For the engineer, as it was for Oliver Heaviside, the primary use of Laplace transforms is the solution of ordinary, constant coefficient, linear differential equations. These equations are important not only because they appear in many engineering problems but also because they may serve as approximations, even if locally, to ordinary differential equations with nonconstant coefficients or to nonlinear ordinary differential equations.

For all of these reasons, we wish to solve the *initial-value problem*

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_{n-1} \frac{dy}{dt} + a_n y = f(t), \quad t > 0, \qquad \mathbf{(12.8.1)}$$

by Laplace transforms, where $a_1, a_2, \ldots$ are constants and we know the value of $y, y', \ldots, y^{(n-1)}$ at $t = 0$. The procedure is as follows. Applying the derivative rule Equation 12.1.18 to Equation 12.8.1, we reduce the *differential* equation to an *algebraic* one involving the constants $a_1, a_2, \ldots, a_n$, the parameter s, the Laplace transform of $f(t)$, and the values of the initial conditions. We then solve for the Laplace transform of $y(t)$, $Y(s)$. Finally, we apply one of the many techniques of inverting a Laplace transform to find $y(t)$.

Similar considerations hold with *systems* of ordinary differential equations. The Laplace transform of the system of ordinary differential equations results in an algebraic set of equations containing $Y_1(s), Y_2(s), \ldots, Y_n(s)$. By some method we solve this set of equations and invert each transform $Y_1(s), Y_2(s), \ldots, Y_n(s)$ in turn to give $y_1(t), y_2(t), \ldots, y_n(t)$.

The following examples will illustrate the details of the process.

• Example 12.8.1

Let us solve the ordinary differential equation

$$y'' + 2y' = 8t, \qquad \mathbf{(12.8.2)}$$

subject to the initial conditions that $y'(0) = y(0) = 0$. Taking the Laplace transform of both sides of Equation 12.8.2,

$$\mathcal{L}(y'') + 2\mathcal{L}(y') = 8\mathcal{L}(t), \tag{12.8.3}$$

or

$$s^2Y(s) - sy(0) - y'(0) + 2sY(s) - 2y(0) = \frac{8}{s^2}, \tag{12.8.4}$$

where $Y(s) = \mathcal{L}[y(t)]$. Substituting the initial conditions into Equation 12.8.4 and solving for $Y(s)$,

$$Y(s) = \frac{8}{s^3(s+2)} = \frac{A}{s^3} + \frac{B}{s^2} + \frac{C}{s} + \frac{D}{s+2} = \frac{(s+2)A + s(s+2)B + s^2(s+2)C + s^3D}{s^3(s+2)}. \tag{12.8.5}$$

Matching powers of s in the numerators of Equation 12.8.5, $C + D = 0$, $B + 2C = 0$, $A + 2B = 0$, and $2A = 8$ or $A = 4$, $B = -2$, $C = 1$, and $D = -1$. Therefore,

$$Y(s) = \frac{4}{s^3} - \frac{2}{s^2} + \frac{1}{s} - \frac{1}{s+2}. \tag{12.8.6}$$

Finally, performing term-by-term inversion of Equation 12.8.6, the final solution is

$$y(t) = 2t^2 - 2t + 1 - e^{-2t}. \tag{12.8.7}$$

We could have done the same operations using the symbolic toolbox with MATLAB. The MATLAB script

```
clear
% define symbolic variables
syms s t Y
% take Laplace transform of left side of differential equation
LHS = laplace(diff(diff(sym('y(t)')))+2*diff(sym('y(t)')));
% take Laplace transform of right side of differential equation
RHS = laplace(8*t);
% set Y for Laplace transform of y
%     and introduce initial conditions
newLHS = subs(LHS,'laplace(y(t),t,s)','y(0)','D(y)(0)',Y,0,0);
% solve for Y
Y = solve(newLHS-RHS,Y);
% invert Laplace transform and find y(t)
y = ilaplace(Y,s,t)
```

yields the result

```
y =
1-exp(-2*t)-2*t+2*t^2
```

which agrees with Equation 12.8.7. □

• Example 12.8.2

Let us solve the ordinary differential equation

$$y'' + y = H(t) - H(t-1) \tag{12.8.8}$$

with the initial conditions that $y'(0) = y(0) = 0$. Taking the Laplace transform of both sides of Equation 12.8.8,

$$s^2Y(s) - sy(0) - y'(0) + Y(s) = \frac{1}{s} - \frac{e^{-s}}{s}, \tag{12.8.9}$$

where $Y(s) = \mathcal{L}[y(t)]$. Substituting the initial conditions into Equation 12.8.9 and solving for $Y(s)$,

$$Y(s) = \left(\frac{1}{s} - \frac{s}{s^2+1}\right) - \left(\frac{1}{s} - \frac{s}{s^2+1}\right)e^{-s}. \tag{12.8.10}$$

Using the second shifting theorem, the final solution is

$$y(t) = 1 - \cos(t) - [1 - \cos(t-1)]H(t-1). \tag{12.8.11}$$

We can check our results using the MATLAB script

```
clear
% define symbolic variables
syms s t Y
% take Laplace transform of left side of differential equation
LHS = laplace(diff(diff(sym('y(t)')))+sym('y(t)'));
% take Laplace transform of right side of differential equation
RHS = laplace('Heaviside(t) - Heaviside(t-1)',t,s);
% set Y for Laplace transform of y
%     and introduce initial conditions
newLHS = subs(LHS,'laplace(y(t),t,s)','y(0)','D(y)(0)',Y,0,0);
% solve for Y
Y = solve(newLHS-RHS,Y);
% invert Laplace transform and find y(t)
y = ilaplace(Y,s,t)
```

which yields

```
y =
1-cos(t)-Heaviside(t-1)+Heaviside(t-1)*cos(t-1)
```

□

• Example 12.8.3

Let us solve the ordinary differential equation

$$y'' + 2y' + y = f(t) \tag{12.8.12}$$

with the initial conditions that $y'(0) = y(0) = 0$, where $f(t)$ is an unknown function whose Laplace transform exists. Taking the Laplace transform of both sides of Equation 12.8.12,

$$s^2Y(s) - sy(0) - y'(0) + 2sY(s) - 2y(0) + Y(s) = F(s), \tag{12.8.13}$$

where $Y(s) = \mathcal{L}[y(t)]$. Substituting the initial conditions into Equation 12.8.13 and solving for $Y(s)$,

$$Y(s) = \frac{1}{(s+1)^2}F(s). \tag{12.8.14}$$

We wrote Equation 12.8.14 in this form because the transform $Y(s)$ equals the product of two transforms $1/(s+1)^2$ and $F(s)$. Therefore, by the convolution theorem we can immediately write

$$y(t) = te^{-t} * f(t) = \int_0^t xe^{-x} f(t-x)\,dx. \tag{12.8.15}$$

Without knowing $f(t)$, this is as far as we can go. □

• Example 12.8.4: Forced harmonic oscillator

Let us solve the *simple harmonic oscillator* forced by a harmonic forcing

$$y'' + \omega^2 y = \cos(\omega t), \tag{12.8.16}$$

subject to the initial conditions that $y'(0) = y(0) = 0$. Although the complete solution could be found by summing the complementary solution and a particular solution obtained, say, from the method of undetermined coefficients, we now illustrate how we can use Laplace transforms to solve this problem.

Taking the Laplace transform of both sides of Example 12.8.16, substituting in the initial conditions, and solving for $Y(s)$,

$$Y(s) = \frac{s}{(s^2+\omega^2)^2}, \tag{12.8.17}$$

and

$$y(t) = \frac{1}{\omega}\sin(\omega t) * \cos(\omega t) = \frac{t}{2\omega}\sin(\omega t). \tag{12.8.18}$$

Equation 12.8.18 gives an oscillation that grows linearly with time although the forcing function is simply periodic. Why does this occur? Recall that our simple harmonic oscillator has the natural frequency ω. But that is exactly the frequency at which we drive the system. Consequently, our choice of forcing has resulted in *resonance* where energy continuously feeds into the oscillator. □

• Example 12.8.5

Let us solve the *system* of ordinary differential equations:

$$2x' + y = \cos(t), \tag{12.8.19}$$

and

$$y' - 2x = \sin(t), \tag{12.8.20}$$

subject to the initial conditions that $x(0) = 0$, and $y(0) = 1$. Taking the Laplace transform of Equation 12.8.19 and Equation 12.8.20,

$$2sX(s) + Y(s) = \frac{s}{s^2+1}, \tag{12.8.21}$$

and

$$-2X(s) + sY(s) = 1 + \frac{1}{s^2+1}, \tag{12.8.22}$$

after introducing the initial conditions. Solving for $X(s)$ and $Y(s)$,

$$X(s) = -\frac{1}{(s^2+1)^2}, \tag{12.8.23}$$

and

$$Y(s) = \frac{s}{s^2+1} + \frac{2s}{(s^2+1)^2}. \tag{12.8.24}$$

Taking the inverse of Equation 12.8.23 and Equation 12.8.24 term by term,

$$x(t) = \tfrac{1}{2}[t\cos(t) - \sin(t)], \tag{12.8.25}$$

and

$$y(t) = t\sin(t) + \cos(t). \tag{12.8.26}$$

The MATLAB script

```
clear
% define symbolic variables
syms s t X Y
% take Laplace transform of left side of differential equations
LHS1 = laplace(2*diff(sym('x(t)'))+sym('y(t)'));
LHS2 = laplace(diff(sym('y(t)'))-2*sym('x(t)'));
% take Laplace transform of right side of differential equations
RHS1 = laplace(cos(t)); RHS2 = laplace(sin(t));
% set X and Y for Laplace transforms of x and y
%     and introduce initial conditions
newLHS1 = subs(LHS1,'laplace(x(t),t,s)','laplace(y(t),t,s)',...
    'x(0)','y(0)',X,Y,0,1);
newLHS2 = subs(LHS2,'laplace(x(t),t,s)','laplace(y(t),t,s)',...
    'x(0)','y(0)',X,Y,0,1);
% solve for X and Y
[X,Y] = solve(newLHS1-RHS1,newLHS2-RHS2,X,Y);
% invert Laplace transform and find x(t) and y(t)
x = ilaplace(X,s,t); y = ilaplace(Y,s,t)
```

uses the symbolic toolbox to solve Equation 12.8.19 and Equation 12.8.20. MATLAB finally gives

```
x =
1/2*t*cos(t)-1/2*sin(t)
y =
t*sin(t)+cos(t)
```

□

• Example 12.8.6

Let us determine the displacement of a mass m attached to a spring and excited by the driving force

$$F(t) = mA\left(1 - \frac{t}{T}\right)e^{-t/T}. \tag{12.8.27}$$

The dynamical equation governing this system is

$$y'' + \omega^2 y = A\left(1 - \frac{t}{T}\right)e^{-t/T}, \tag{12.8.28}$$

where $\omega^2 = k/m$ and k is the spring constant. Assuming that the system is initially at rest, the Laplace transform of the dynamical system is

$$(s^2 + \omega^2)Y(s) = \frac{A}{s + 1/T} - \frac{A}{T(s + 1/T)^2}, \tag{12.8.29}$$

or

$$Y(s) = \frac{A}{(s^2 + \omega^2)(s + 1/T)} - \frac{A}{T(s^2 + \omega^2)(s + 1/T)^2}. \tag{12.8.30}$$

Partial fractions yield

$$\begin{aligned} Y(s) = \frac{A}{\omega^2 + 1/T^2}\left(\frac{1}{s + 1/T} - \frac{s - 1/T}{s^2 + \omega^2}\right) - \frac{A}{T(\omega^2 + 1/T^2)^2} \\ \times \left[\frac{1/T^2 - \omega^2}{s^2 + \omega^2} - \frac{2s/T}{s^2 + \omega^2} + \frac{\omega^2 + 1/T^2}{(s + 1/T)^2} + \frac{2/T}{s + 1/T}\right]. \end{aligned} \tag{12.8.31}$$

Inverting Equation 12.8.31 term by term,

$$\begin{aligned} y(t) = \frac{AT^2}{1 + \omega^2 T^2}\left[e^{-t/T} - \cos(\omega t) + \frac{\sin(\omega t)}{\omega T}\right] - \frac{AT^2}{(1 + \omega^2 T^2)^2}\bigg\{(1 - \omega^2 T^2)\frac{\sin(\omega t)}{\omega T} \\ + 2\left[e^{-t/T} - \cos(\omega t)\right] + (1 + \omega^2 T^2)(t/T)e^{-t/T}\bigg\}. \end{aligned} \tag{12.8.32}$$

The solution to this problem consists of two parts. The exponential terms result from the forcing and will die away with time. This is the *transient* portion of the solution. The sinusoidal terms are those natural oscillations that are necessary so that the solution satisfies the initial conditions. They are the *steady-state* portion of the solution and endure forever. Figure 12.8.1 illustrates the solution when $\omega T = 0.1$, 1, and 2. Note that the displacement decreases in magnitude as the nondimensional frequency of the oscillator increases. □

• Example 12.8.7

Let us solve the equation

$$y'' + 16y = \delta(t - \pi/4) \tag{12.8.33}$$

with the initial conditions that $y(0) = 1$, and $y'(0) = 0$.

Taking the Laplace transform of Equation 12.8.33 and inserting the initial conditions,

$$(s^2 + 16)Y(s) = s + e^{-s\pi/4}, \tag{12.8.34}$$

or

$$Y(s) = \frac{s}{s^2 + 16} + \frac{e^{-s\pi/4}}{s^2 + 16}. \tag{12.8.35}$$

Applying the second shifting theorem,

$$y(t) = \cos(4t) + \tfrac{1}{4}\sin[4(t - \pi/4)]H(t - \pi/4) = \cos(4t) - \tfrac{1}{4}\sin(4t)H(t - \pi/4). \tag{12.8.36}$$

We can check our results using the MATLAB script

```
clear
% define symbolic variables
syms pi s t Y
```

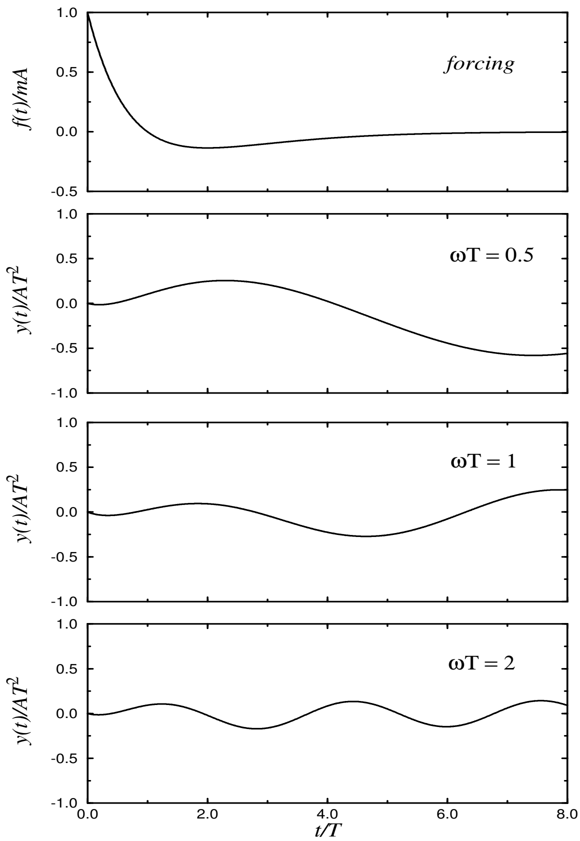


Figure 12.8.1: Displacement of a simple harmonic oscillator with nondimensional frequency ωT as a function of time t/T. The top frame shows the forcing function.

```
% take Laplace transform of left side of differential equation
LHS = laplace(diff(diff(sym('y(t)')))+16*sym('y(t)'));
% take Laplace transform of right side of differential equation
RHS = laplace('Dirac(t-pi/4)',t,s);
% set Y for Laplace transform of y
%     and introduce initial conditions
newLHS = subs(LHS,'laplace(y(t),t,s)','y(0)','D(y)(0)',Y,1,0);
% solve for Y
Y = solve(newLHS-RHS,Y);
% invert Laplace transform and find y(t)
y = ilaplace(Y,s,t)
```

which yields

```
y =
cos(4*t)-1/4*Heaviside(t-1/4*pi)*sin(4*t)
```

We can also verify that Equation 12.8.36 is the solution to our initial-value problem by computing the (generalized) derivative of Equation 12.8.36, or

$$y'(t) = -4\sin(4t) - \cos(4t)H(t-\pi/4) - \tfrac{1}{4}\sin(4t)\delta(t-\pi/4) \quad \textbf{(12.8.37)}$$

$$= -4\sin(4t) - \cos(4t)H(t-\pi/4) - \tfrac{1}{4}\sin(\pi)\delta(t-\pi/4) \quad \textbf{(12.8.38)}$$

$$= -4\sin(4t) - \cos(4t)H(t-\pi/4), \quad \textbf{(12.8.39)}$$

since $f(t)\delta(t-t_0) = f(t_0)\delta(t-t_0)$. Similarly,

$$y''(t) = -16\cos(4t) + 4\sin(4t)H(t-\pi/4) - \cos(4t)\delta(t-\pi/4) \tag{12.8.40}$$

$$= -16\cos(4t) + 4\sin(4t)H(t-\pi/4) - \cos(\pi)\delta(t-\pi/4) \tag{12.8.41}$$

$$= -16\cos(4t) + 4\sin(4t)H(t-\pi/4) + \delta(t-\pi/4). \tag{12.8.42}$$

Substituting Equation 12.8.36 and Equation 12.8.42 into Equation 12.8.32 completes the verification. A quick check of $y(0)$ and $y'(0)$ also shows that we have the correct solution.□

• **Example 12.8.8: Oscillations in electric circuits**

During the middle of the nineteenth century, Lord Kelvin[19] analyzed the LCR electrical circuit shown in Figure 12.8.2, which contains resistance R, capacitance C, and inductance L. For reasons that we shall shortly show, this LCR circuit has become one of the quintessential circuits for electrical engineers. In this example, we shall solve the problem by Laplace transforms.

Because we can add the potential differences across the elements, the equation governing the LCR circuit is

$$L\frac{dI}{dt} + RI + \frac{1}{C}\int_0^t I\,d\tau = E(t), \tag{12.8.43}$$

where I denotes the current in the circuit. Let us solve Equation 12.8.43 when we close the circuit and the initial conditions are $I(0) = 0$ and $Q(0) = -Q_0$. Taking the Laplace transform of Equation 12.8.43,

$$\left(Ls + R + \frac{1}{Cs}\right)\overline{I}(s) = LI(0) - \frac{Q(0)}{Cs}. \tag{12.8.44}$$

Solving for $\overline{I}(s)$,

$$\overline{I}(s) = \frac{Q_0}{Cs(Ls+R+1/Cs)} = \frac{\omega_0^2 Q_0}{s^2+2\alpha s+\omega_0^2} = \frac{\omega_0^2 Q_0}{(s+\alpha)^2+\omega_0^2-\alpha^2}, \tag{12.8.45}$$

where $\alpha = R/(2L)$, and $\omega_0^2 = 1/(LC)$. From the first shifting theorem,

$$I(t) = \frac{\omega_0^2 Q_0}{\omega} e^{-\alpha t}\sin(\omega t), \tag{12.8.46}$$

where $\omega^2 = \omega_0^2 - \alpha^2 > 0$. The quantity ω is the natural frequency of the circuit, which is lower than the free frequency ω_0 of a circuit formed by a condenser and coil. Most importantly, the solution decays in amplitude with time.

Although Kelvin's solution was of academic interest when he originally published it, this radically changed with the advent of radio telegraphy[20] because the LCR circuit described the fundamental physical properties of wireless transmitters and receivers.[21] The

---

[19] Thomson, W., 1853: On transient electric currents. *Philos. Mag., Ser. 4*, **5**, 393–405.

[20] Stone, J S., 1914: The resistance of the spark and its effect on the oscillations of electrical oscillators. *Proc. IRE*, **2**, 307–324.

[21] See Hogan, J. L., 1916: Physical aspects of radio telegraphy. *Proc. IRE*, **4**, 397–420.

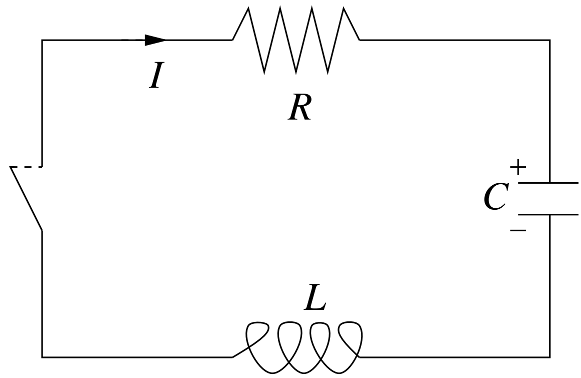


Figure 12.8.2: Schematic of a LCR circuit.

inescapable conclusion from numerous analyses was that no matter how cleverly the receiver was designed, eventually the resistance in the circuit would dampen the electrical oscillations and thus limit the strength of the received signal.

This technical problem was overcome by Armstrong,[22] who invented an electrical circuit that used De Forest's audion (the first vacuum tube) for generating electrical oscillations and for amplifying externally impressed oscillations by "regenerative action." The effect of adding the "thermionic amplifier" is seen by again considering the LRC circuit as shown in Figure 12.8.3 with the modification suggested by Armstrong.[23]

The governing equations of this new circuit are

$$L_1 \frac{dI}{dt} + RI + \frac{1}{C} \int_0^t I \, d\tau + M \frac{dI_p}{dt} = 0, \tag{12.8.47}$$

and

$$L_2 \frac{dI_p}{dt} + R_0 I_p + M \frac{dI}{dt} + \frac{\mu}{C} \int_0^t I \, d\tau = 0, \tag{12.8.48}$$

where the plate circuit has the current I_p, the resistance R_0, the inductance L_2, and the electromotive force (emf) of $\mu \int_0^t I \, d\tau / C$. The mutual inductance between the two circuits is given by M. Taking the Laplace transform of Equation 12.8.47 and Equation 12.8.48,

$$L_1 s \overline{I}(s) + R\overline{I}(s) + \frac{\overline{I}(s)}{sC} + Ms\overline{I}_p(s) = \frac{Q_0}{sC}, \tag{12.8.49}$$

and

$$L_2 s \overline{I}_p(s) + R_0 \overline{I}_p(s) + Ms\overline{I}(s) + \frac{\mu}{sC}\overline{I}(s) = 0. \tag{12.8.50}$$

Eliminating $\overline{I}_p(s)$ between Equation 12.8.49 and Equation 12.8.50 and solving for $\overline{I}(s)$,

$$\overline{I}(s) = \frac{(L_2 s + R_0)Q_0}{(L_1 L_2 - M^2)Cs^3 + (RL_2 + R_0 L_1)Cs^2 + (L_2 + CRR_0 - \mu M)s + R_0}. \tag{12.8.51}$$

[22] Armstrong, E. H., 1915: Some recent developments in the audion receiver. *Proc. IRE*, **3**, 215–247.

[23] See Ballantine, S., 1919: The operational characteristics of thermionic amplifiers. *Proc. IRE*, **7**, 129–161.

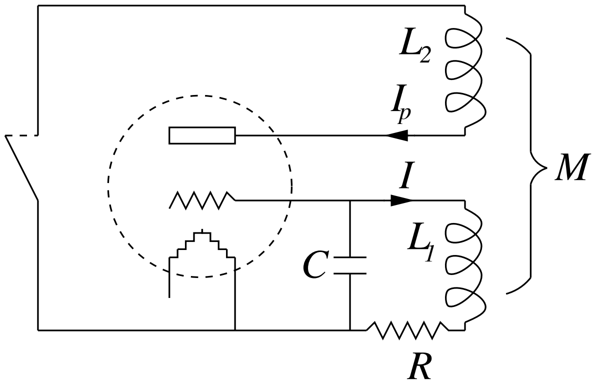


Figure 12.8.3: Schematic of an LCR circuit with the addition of a thermionic amplifier. (From Ballantine, S., 1919: The operational characteristics of thermionic amplifiers. *Proc. IRE*, **7**, 155.)

For high-frequency radio circuits, we can approximate the roots of the denominator of Equation 12.8.51 as

$$s_1 \approx -\frac{R_0}{L_2 + CRR_0 - \mu M}, \qquad (\mathbf{12.8.52})$$

and

$$s_{2,3} \approx \frac{R_0}{2(L_2 + CRR_0 - \mu M)} - \frac{R_0 L_1 + RL_2}{2(L_1 L_2 - M^2)} \pm i\omega. \qquad (\mathbf{12.8.53})$$

In the limit of M and R_0 vanishing, we recover our previous result for the LRC circuit. However, in reality, R_0 is very large and our solution has three terms. The term associated with s_1 is a rapidly decaying transient while the s_2 and s_3 roots yield oscillatory solutions with a *slight* amount of damping. Thus, our analysis shows that in the ordinary regenerative circuit, the tube effectively introduces sufficient "negative" resistance so that the resultant positive resistance of the equivalent LCR circuit is relatively low, and the response of an applied signal voltage at the resonant frequency of the circuit is therefore relatively great. Later, Armstrong[24] extended his work on regeneration by introducing an electrical circuit - the superregenerative circuit - where the regeneration is made large enough so that the resultant resistance is negative, and self-sustained oscillations can occur.[25] It was this circuit[26] that led to the explosive development of radio in the 1920s and 1930s. □

• Example 12.8.9: Resonance transformer circuit

One of the fundamental electrical circuits of early radio telegraphy[27] is the resonance transformer circuit shown in Figure 12.8.4. Its development gave transmitters and receivers the ability to tune to each other.

The governing equations follow from Kirchhoff's law and are

$$L_1 \frac{dI_1}{dt} + M \frac{dI_2}{dt} + \frac{1}{C_1} \int_0^t I_1 \, d\tau = E(t), \qquad (\mathbf{12.8.54})$$

[24] Armstrong, E. H., 1922: Some recent developments of regenerative circuits. *Proc. IRE*, **10**, 244–260.

[25] See Frink, F. W., 1938: The basic principles of superregenerative reception. *Proc. IRE*, **26**, 76–106.

[26] Lewis, T., 1991: *Empire of the Air: The Men Who Made Radio*. HarperCollins Publishers, 421 pp.

[27] Fleming, J. A., 1919: *The Principles of Electric Wave Telegraphy and Telephony*. Longmans, Green, 911 pp.

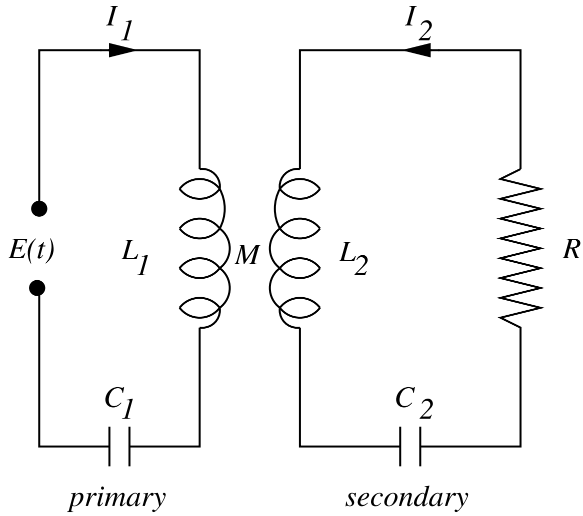


Figure 12.8.4: Schematic of a resonance transformer circuit.

and

$$M\frac{dI_1}{dt} + L_2\frac{dI_2}{dt} + RI_2 + \frac{1}{C_2}\int_0^t I_2\,d\tau = 0. \tag{12.8.55}$$

Let us examine the oscillations generated if initially the system has no currents or charges and the forcing function is $E(t) = \delta(t)$.

Taking the Laplace transform of Equation 12.8.54 and Equation 12.8.55,

$$L_1 s\overline{I}_1 + Ms\overline{I}_2 + \frac{\overline{I}_1}{sC_1} = 1, \tag{12.8.56}$$

and

$$Ms\overline{I}_1 + L_2 s\overline{I}_2 + R\overline{I}_2 + \frac{\overline{I}_2}{sC_2} = 0. \tag{12.8.57}$$

Because the current in the second circuit is of greater interest, we solve for $\overline{I}_2$ and find that

$$\overline{I}_2(s) = -\frac{Ms^3}{L_1L_2[(1-k^2)s^4 + 2\alpha\omega_2^2 s^3 + (\omega_1^2+\omega_2^2)s^2 + 2\alpha\omega_1^2 s + \omega_1^2\omega_2^2]}, \tag{12.8.58}$$

where $\alpha = R/(2L_2)$, $\omega_1^2 = 1/(L_1C_1)$, $\omega_2^2 = 1/(L_2C_2)$, and $k^2 = M^2/\ (L_1L_2)$, the so-called coefficient of coupling.

We can obtain analytic solutions if we assume that the coupling is weak ($k^2 \ll 1$). Equation 12.8.58 becomes

$$\overline{I}_2 = -\frac{Ms^3}{L_1L_2(s^2+\omega_1^2)(s^2+2\alpha s+\omega_2^2)}. \tag{12.8.59}$$

Using partial fractions and inverting term by term, we find that

$$I_2(t) = \frac{M}{L_1L_2}\left[\frac{2\alpha\omega_1^3\sin(\omega_1 t)}{(\omega_2^2-\omega_1^2)^2+4\alpha^2\omega_1^2} + \frac{\omega_1^2(\omega_2^2-\omega_1^2)\cos(\omega_1 t)}{(\omega_2^2-\omega_1^2)^2+4\alpha^2\omega_1^2}\right. \tag{12.8.60}$$
$$\left. + \frac{\alpha\omega_2^4 - 3\alpha\omega_1^2\omega_2^2 + 4\alpha^3\omega_1^2}{(\omega_2^2-\omega_1^2)^2+4\alpha^2\omega_1^2}e^{-\alpha t}\frac{\sin(\omega t)}{\omega} - \frac{\omega_2^2(\omega_2^2-\omega_1^2)+4\alpha^2\omega_1^2}{(\omega_2^2-\omega_1^2)^2+4\alpha^2\omega_1^2}e^{-\alpha t}\cos(\omega t)\right],$$

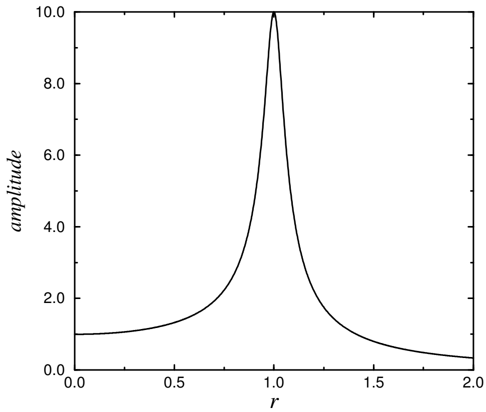


Figure 12.8.5: The resonance curve $1/\sqrt{(r^2-1)^2+0.01}$ for a resonance transformer circuit with $r = \omega_2/\omega_1$.

where $\omega^2 = \omega_2^2 - \alpha^2$.

The exponentially damped solutions will eventually disappear, leaving only the steady-state oscillations that vibrate with the angular frequency ω_1, the natural frequency of the primary circuit. If we rewrite this steady-state solution in amplitude/phase form, the amplitude is

$$\frac{M}{L_1L_2\sqrt{(r^2-1)^2+4\alpha^2/\omega_1^2}}, \tag{12.8.61}$$

where $r = \omega_2/\omega_1$. As Figure 12.8.5 shows, as r increases from zero to two, the amplitude rises until a very sharp peak occurs at $r = 1$ and then decreases just as rapidly as we approach $r = 2$. Thus, the resonance transformer circuit provides a convenient way to tune a transmitter or receiver to the frequency ω_1. □

• Example 12.8.10: Delay differential equation

Laplace transforms provide a valuable tool in solving a general class of ordinary differential equations called *delay differential equations*. These equations arise in such diverse fields as chemical kinetics[28] and population dynamics.[29]

To illustrate the technique,[30] consider the differential equation

$$x' = -ax(t-1) \tag{12.8.62}$$

with $x(t) = 1 - at$ for $0 < t < 1$. Clearly, $x(0) = 1$.

Multiplying Equation 12.8.62 by e^{-st} and integrating from 1 to ∞,

$$\int_1^\infty x'(t)e^{-st}\,dt = -a\int_1^\infty x(t-1)e^{-st}\,dt \tag{12.8.63}$$

[28] See Roussel, M. R., 1996: The use of delay differential equations in chemical kinetics. *J. Phys. Chem.*, **100**, 8323–8330.

[29] See the first chapter of MacDonald, N., 1989: *Biological Delay Systems: Linear Stability Theory*. Cambridge University Press, 235 pp.

[30] See Epstein, I. R., 1990: Differential delay equations in chemical kinetics: Some simple linear model systems. *J. Chem. Phys.*, **92**, 1702–1712.

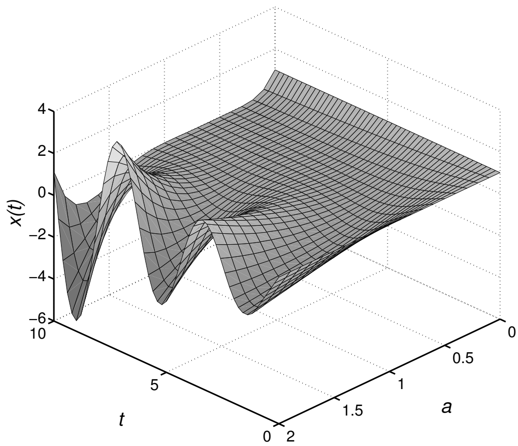


Figure 12.8.6: The solution to the delay differential equation, Equation 12.8.62, at various times t and values of a.

$$\int_0^\infty x'(t)e^{-st}\,dt - \int_0^1 x'(t)e^{-st}\,dt = -a\int_0^\infty x(\tau)e^{-s(\tau+1)}\,d\tau \tag{12.8.64}$$

$$sX(s) - 1 + a\int_0^1 e^{-st}\,dt = -ae^{-s}X(s) \tag{12.8.65}$$

$$sX(s) - 1 - \frac{a}{s}\,e^{-st}\Big|_0^1 = -ae^{-s}X(s) \tag{12.8.66}$$

since $x'(t) = -a$ for $0 < t < 1$. Solving for $X(s)$,

$$X(s) = (1 + ae^{-s}/s - a/s)/[s(1 + ae^{-s}/s)]. \tag{12.8.67}$$

To facilitate the inversion of Equation 12.8.67, we expand its denominator in terms of a geometric series and find that

$$X(s) = \sum_{n=0}^\infty (-a)^n e^{-ns}/s^{n+1} + \sum_{n=0}^\infty (-a)^{n+1}e^{-ns}/s^{n+2} - \sum_{n=0}^\infty (-a)^{n+1}e^{-(n+1)s}/s^{n+2}. \tag{12.8.68}$$

The first and third sums cancel, except for the $n = 0$ term in the first sum. Therefore,

$$X(s) = \frac{1}{s} + \sum_{n=0}^\infty (-a)^{n+1}e^{-ns}/s^{n+2} \tag{12.8.69}$$

and

$$x(t) = 1 + \sum_{n=0}^\infty \frac{(-a)^{n+1}}{(n+1)!}H(t-n)(t-n)^{n+1}. \tag{12.8.70}$$

Figure 12.8.6 illustrates Equation 12.8.70 as a function of time for various values of a. For $0 < a < e^{-1}$, $x(t)$ decays monotonically from 1 to an asymptotic limit of zero. For $e^{-1} < a < \pi/2$, the solution is a damped oscillatory function. If $\pi/2 < a$, then $x(t)$ is oscillatory with an exponentially increasing envelope. When $a = \pi/2$, $x(t)$ oscillates periodically. □

• Example 12.8.11

Laplace transforms can sometimes be used to solve ordinary differential equations where the coefficients are powers of t. To illustrate this, let us solve

$$y'' + 2ty' - 4y = 0, \qquad y(0) = 1, \quad \lim_{t\to\infty} y(t) \to 0. \tag{12.8.71}$$

We begin by taking the Laplace transform of Equation 12.8.71 and find that

$$s^2 Y(s) - sy(0) - y'(0) - 2\frac{d}{ds}[sY(s) - y(0)] - 4Y(s) = 0. \tag{12.8.72}$$

An interesting aspect of this problem is the fact that we do not know $y'(0)$. To circumvent this difficulty, let us temporarily set $y'(0) = -A$ so that Equation 12.8.72 becomes

$$\frac{dY}{ds} + \left(\frac{3}{s} - \frac{s}{2}\right) Y = \frac{A}{2s} - \frac{1}{2}. \tag{12.8.73}$$

Later on, we will find A.

Equation 12.8.73 is a first-order, linear, ordinary differential equation with s as its independent variable. To find $Y(s)$, we use the standard technique of multiplying it by its integrating factor, here $\mu(s) = s^3 e^{-s^2/4}$, and rewriting it as

$$\frac{d}{ds}\left[s^3 e^{-s^2/4} Y(s)\right] = \tfrac{1}{2} A s^2 e^{-s^2/4} - \tfrac{1}{2} s^3 e^{-s^2/4}. \tag{12.8.74}$$

Integrating Equation 12.8.74 from s to ∞, we obtain

$$s^3 e^{-s^2/4} Y(s) = (s^2 + 4)e^{-s^2/4} - A\left[se^{-s^2/4} + \sqrt{\pi}\,\mathrm{erfc}(s/2)\right], \tag{12.8.75}$$

or

$$Y(s) = \frac{4}{s^3} + \frac{1}{s} - \frac{A}{s^2} - \frac{A\sqrt{\pi}}{s^3} e^{s^2/4}\,\mathrm{erfc}(s/2). \tag{12.8.76}$$

We must now evaluate A. From the final-value theorem, $\lim_{t\to\infty} y(t) = \lim_{s\to 0} sY(s) = 0$. Therefore, multiplying Equation 12.8.76 by s and using the expansion for the complementary error function for small s, we have that

$$sY(s) = \frac{4}{s^2} + 1 - \frac{A}{s} - \frac{A\sqrt{\pi}}{s^2}\left[1 + \frac{s^2}{4} - \frac{s}{\sqrt{\pi}} + \cdots\right]. \tag{12.8.77}$$

In order that $\lim_{s\to 0} sY(s) = 0$, $A = 4/\sqrt{\pi}$. Therefore,

$$Y(s) = \frac{4}{s^3} + \frac{1}{s} - \frac{4}{\sqrt{\pi}\, s^2} - \frac{4}{s^3} e^{s^2/4}\,\mathrm{erfc}(s/2). \tag{12.8.78}$$

The final step is to invert Equation 12.8.78. Applying tables and the convolution theorem,

$$y(t) = 2t^2 + 1 - \frac{4t}{\sqrt{\pi}} - \frac{4}{\sqrt{\pi}} \int_0^t (t-x)^2 e^{-x^2}\, dx = (2t^2 + 1)[1 - \mathrm{erf}(t)] - \frac{2t}{\sqrt{\pi}} e^{-t^2}. \tag{12.8.79}$$

Problems

Solve the following ordinary differential equations by Laplace transforms. Then use MATLAB to verify your solution.

1. $y' - 2y = 1 - t; \quad y(0) = 1$

2. $y'' - 4y' + 3y = e^t; \quad y(0) = 0, y'(0) = 0$

3. $y'' - 4y' + 3y = e^{2t}; \quad y(0) = 0, y'(0) = 1$

4. $y'' - 6y' + 8y = e^t; \quad y(0) = 3, y'(0) = 9$

5. $y'' + 4y' + 3y = e^{-t}; \quad y(0) = 1, y'(0) = 1$

6. $y'' + y = t; \quad y(0) = 1, y'(0) = 0$

7. $y'' + 4y' + 3y = e^t; \quad y(0) = 0, y'(0) = 2$

8. $y'' - 4y' + 5y = 0; \quad y(0) = 2, y'(0) = 4$

9. $y' + y = tH(t-1); \quad y(0) = 0$

10. $y'' + 3y' + 2y = H(t-1); \quad y(0) = 0, y'(0) = 1$

11. $y'' - 3y' + 2y = H(t-1); \quad y(0) = 0, y'(0) = 1$

12. $y'' + 4y = 3H(t-4); \quad y(0) = 1, y'(0) = 0$

13. $y'' + 4y' + 4y = 4H(t-2); \quad y(0) = 0, y'(0) = 0$

14. $y'' + 3y' + 2y = e^{t-1}H(t-1); \quad y(0) = 0, y'(0) = 1$

15. $y'' - 3y' + 2y = e^{-(t-2)}H(t-2); \quad y(0) = 0, y'(0) = 0$

16. $y'' - 3y' + 2y = H(t-1) - H(t-2); \quad y(0) = 0, y'(0) = 0$

17. $y'' + y = 1 - H(t-T); \quad y(0) = 0, y'(0) = 0$

18. $y'' + y = \begin{cases} \sin(t), & 0 \le t \le \pi, \\ 0, & \pi \le t; \end{cases} \quad y(0) = 0, y'(0) = 0$

19. $y'' + 3y' + 2y = \begin{cases} t, & 0 \le t \le a, \\ ae^{-(t-a)}, & a \le t; \end{cases} \quad y(0) = 0, y'(0) = 0$

20. $y'' + \omega^2 y = \begin{cases} t/a, & 0 \le t \le a, \\ 1 - (t-a)/(b-a), & a \le t \le b, \\ 0, & b \le t; \end{cases} \quad y(0) = 0, y'(0) = 0$

21. $y'' - 2y' + y = 3\delta(t-2); \quad y(0) = 0, y'(0) = 1$

22. $y'' - 5y' + 4y = \delta(t-1); \quad y(0) = 0, y'(0) = 0$

23. $y'' + 5y' + 6y = 3\delta(t-2) - 4\delta(t-5); \quad y(0) = y'(0) = 0$

24. $y'' + \omega y' = A\delta(t-\tau) - BH(t-\tau); \quad y(0) = y'(0) = 0$

25. $x' - 2x + y = 0, \quad y' - 3x - 4y = 0; \quad x(0) = 1, y(0) = 0$

26. $x' - 2y' = 1, \quad x' + y - x = 0; \quad x(0) = y(0) = 0$

27. $x' + 2x - y' = 0, \quad x' + y + x = t^2; \quad x(0) = y(0) = 0$

28. $x' + 3x - y = 1, \quad x' + y' + 3x = 0; \quad x(0) = 2, y(0) = 0$

29. Forster, Escobal, and Lieske[31] used Laplace transforms to solve the linearized equations of motion of a vehicle in a gravitational field created by two other bodies. A simplified form of this problem involves solving the following system of ordinary differential equations:

$$x'' - 2y' = F_1 + x + 2y, \qquad 2x' + y'' = F_2 + 2x + 3y,$$

subject to the initial conditions that $x(0) = y(0) = x'(0) = y'(0) = 0$. Find the solution to this system.

Use Laplace transforms to find the solution for the following ordinary differential equations:

30. $y'' + 2ty' - 8y = 0, \qquad y(0) = 1, \quad y'(0) = 0$
Step 1: Show that the Laplace transform for this differential equation is $2sY'(s) + (10 - s^2)Y(s) = -s$.
Step 2: Solve these first-order ordinary differential equations and show that $Y(s) = 1/s + 8/s^3 + 32/s^5 + Ae^{s^2/4}/s^5$.
Step 3: Invert $Y(s)$ and show that the general solution is $y(t) = 1 + 4t^2 + 4t^4/3$.

31. $y'' - ty' + 2y = 0, \qquad y(0) = -1, \quad y'(0) = 0$
Step 1: Show that the Laplace transform for this differential equation is $sY'(s) + (s^2 + 3)Y(s) = -s$.
Step 2: Solve the first-order ordinary differential equations and show that $Y(s) = (A - 2)e^{-s^2/2}/s^3 + 2/s^3 - 1/s$.
Step 3: Invert $Y(s)$ and show that the general solution is $y(t) = t^2 - 1$.

32. $ty'' - (2-t)y' - y = 0$
Step 1: Show that the Laplace transform for this differential equation is $s(s+1)Y'(s) + 2(2s+1)Y(s) = 3y(0)$.
Step 2: Solve the first-order ordinary differential equation and show that $Y(s) = y(0)/(s+1) + y(0)/[2(s+1)^2] + A/[s^2(s+1)^2]$.
Step 3: Invert $Y(s)$ and show that the general solution is $y(t) = C_1(t+2)e^{-t} + C_2(t-2)$.

---

[31] Forster, K., P. R. Escobal, and H. A. Lieske, 1968: Motion of a vehicle in the transition region of the three-body problem. *Astronaut. Acta*, **14**, 1–10.

33. $ty'' - 2(a + bt)y' + b(2a + bt)y = 0, \qquad a \geq 0$

Step 1: Show that the Laplace transform for this differential equation is $(s-b)^2 Y'(s) + 2(1+a)(s-b)Y(s) = (1+2a)y(0)$.

Step 2: Solve the first-order ordinary differential equation and show that $Y(s) = y(0)/(s-b) + A/(s-b)^{2+2a}$.

Step 3: Invert $Y(s)$ and show that the general solution is $y(t) = C_1 e^{bt} + C_2 t^{2a+1} e^{bt}$.

12.9 INVERSION BY CONTOUR INTEGRATION

In Section 12.5 and Section 12.6 we showed how we can use partial fractions and convolution to find the inverse of the Laplace transform $F(s)$. In many instances these methods fail simply because of the complexity of the transform to be inverted. In this section we shall show how we can invert transforms through the powerful method of contour integration. Of course, the student must be proficient in the use of complex variables.

Consider the piece-wise differentiable function $f(x)$, which vanishes for $x < 0$. We can express the function $e^{-cx} f(x)$ by the complex Fourier representation of

$$f(x)e^{-cx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega x} \left[\int_0^{\infty} e^{-ct} f(t) e^{-i\omega t} \, dt \right] d\omega, \tag{12.9.1}$$

for any value of the real constant c, where the integral

$$I = \int_0^{\infty} e^{-ct} |f(t)| \, dt \tag{12.9.2}$$

exists. By multiplying both sides of Equation 12.9.1 by e^{cx} and bringing it inside the first integral,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(c+\omega i)x} \left[\int_0^{\infty} f(t) e^{-(c+\omega i)t} \, dt \right] d\omega. \tag{12.9.3}$$

With the substitution $z = c + \omega i$, where z is a new, complex variable of integration,

$$f(x) = \frac{1}{2\pi i} \int_{c-\infty i}^{c+\infty i} e^{zx} \left[\int_0^{\infty} f(t) e^{-zt} \, dt \right] dz. \tag{12.9.4}$$

The quantity inside the square brackets is the Laplace transform $F(z)$. Therefore, we can express $f(t)$ in terms of its transform by the complex contour integral

$$\boxed{f(t) = \frac{1}{2\pi i} \int_{c-\infty i}^{c+\infty i} F(z) e^{tz} dz.} \tag{12.9.5}$$

This line integral, the *Bromwich integral*,[32] runs along the line $x = c$ parallel to the imaginary axis and c units to the right of it, the so-called *Bromwich contour*. We select the value

[32] Bromwich, T. J. I'A., 1916: Normal coordinates in dynamical systems. *Proc. London Math. Soc., Ser. 2*, **15**, 401–448.

An outstanding mathematician at Cambridge University at the turn of the twentieth century, Thomas John I'Anson Bromwich (1875–1929) came to Heaviside's operational calculus through his interest in divergent series. Beginning a correspondence with Heaviside, Bromwich was able to justify operational calculus through the use of contour integrals by 1915. After his premature death, individuals such as J. R. Carson and Sir H. Jeffreys brought Laplace transforms to the increasing attention of scientists and engineers. (Portrait courtesy of the Royal Society of London.)

of c sufficiently large so that the integral, Equation 12.9.2, exists; subsequent analysis shows that this occurs when c is larger than the real part of any of the singularities of $F(z)$.

We must now evaluate the contour integral. Because of the power of the *residue* theorem in complex variables, the contour integral is usually transformed into a closed contour through the use of *Jordan's lemma*. See Section 11.4, Equations 11.4.9 and Equation 11.4.10. The following examples will illustrate the proper use of Equation 12.9.5.

• Example 12.9.1

Let us invert

$$F(s) = \frac{e^{-3s}}{s^2(s-1)}. \tag{12.9.6}$$

From Bromwich's integral,

$$f(t) = \frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} \frac{e^{(t-3)z}}{z^2(z-1)}\,dz = \frac{1}{2\pi i}\oint_C \frac{e^{(t-3)z}}{z^2(z-1)}\,dz - \frac{1}{2\pi i}\int_{C_R} \frac{e^{(t-3)z}}{z^2(z-1)}\,dz, \tag{12.9.7}$$

where C_R is a semicircle of infinite radius in either the right or left half of the z-plane and C is the closed contour that includes C_R and Bromwich's contour. See Figure 12.9.1.

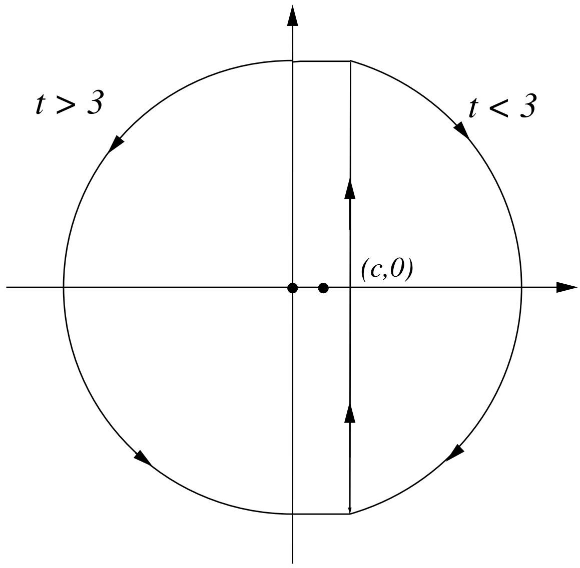


Figure 12.9.1: Contours used in the inversion of Equation 12.9.6.

Our first task is to choose an appropriate contour so that the integral along C_R vanishes. By Jordan's lemma this requires a semicircle in the right half-plane if $t - 3 < 0$ and a semicircle in the left half-plane if $t - 3 > 0$. Consequently, by considering these two separate cases, we force the second integral in Equation 12.9.7 to zero and the inversion simply equals the closed contour.

Consider the case $t < 3$ first. Because Bromwich's contour lies to the right of any singularities, there are no singularities within the closed contour and $f(t) = 0$.

Consider now the case $t > 3$. Within the closed contour in the left half-plane, there is a second-order pole at $z = 0$ and a simple pole at $z = 1$. Therefore,

$$f(t) = \operatorname{Res}\left[\frac{e^{(t-3)z}}{z^2(z-1)}; 0\right] + \operatorname{Res}\left[\frac{e^{(t-3)z}}{z^2(z-1)}; 1\right], \qquad (\mathbf{12.9.8})$$

where

$$\operatorname{Res}\left[\frac{e^{(t-3)z}}{z^2(z-1)}; 0\right] = \lim_{z\to 0} \frac{d}{dz}\left[z^2 \frac{e^{(t-3)z}}{z^2(z-1)}\right] = \lim_{z\to 0} \left[\frac{(t-3)e^{(t-3)z}}{z-1} - \frac{e^{(t-3)z}}{(z-1)^2}\right] = 2 - t, \qquad (\mathbf{12.9.9})$$

and

$$\operatorname{Res}\left[\frac{e^{(t-3)z}}{z^2(z-1)}; 1\right] = \lim_{z\to 1} (z-1)\frac{e^{(t-3)z}}{z^2(z-1)} = e^{t-3}. \qquad (\mathbf{12.9.10})$$

Taking our earlier results into account, the inverse equals

$$f(t) = \left[e^{t-3} - (t-3) - 1\right] H(t-3), \qquad (\mathbf{12.9.11})$$

which we would have obtained from the second shifting theorem and tables. □

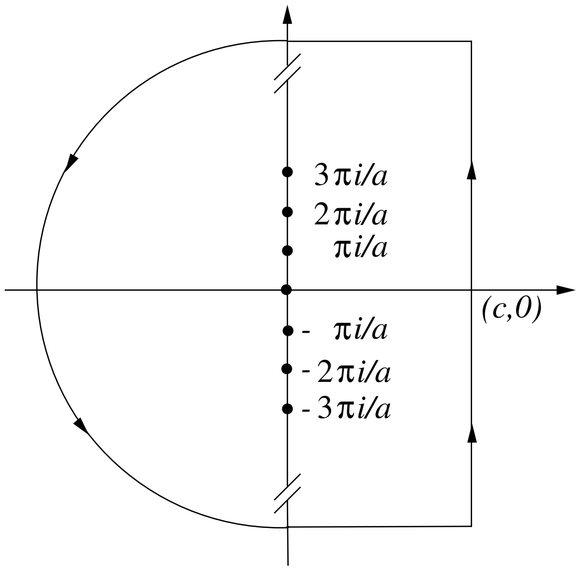


Figure 12.9.2: Contours used in the inversion of Equation 12.9.12.

• Example 12.9.2

For our second example of the inversion of Laplace transforms by complex integration, let us find the inverse of

$$F(s) = \frac{1}{s\sinh(as)}, \tag{12.9.12}$$

where a is real. From Bromwich's integral,

$$f(t) = \frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} \frac{e^{tz}}{z\sinh(az)}\,dz. \tag{12.9.13}$$

Here c is greater than the real part of any of the singularities in Equation 12.9.12. Using the infinite product for the hyperbolic sine,[33]

$$\frac{e^{tz}}{z\sinh(az)} = \frac{e^{tz}}{az^2[1+a^2z^2/\pi^2][1+a^2z^2/(4\pi^2)][1+a^2z^2/(9\pi^2)]\cdots}. \tag{12.9.14}$$

Thus, we have a second-order pole at $z = 0$ and simple poles at $z_n = \pm n\pi i/a$, where $n = 1, 2, 3, \ldots$.

We can convert the line integral Equation 12.9.13, with the Bromwich contour lying parallel and slightly to the right of the imaginary axis, into a closed contour using Jordan's lemma through the addition of an infinite semicircle joining $i\infty$ to $-i\infty$, as shown in Figure 12.9.2. We now apply the residue theorem. For the second-order pole at $z = 0$,

$$\operatorname{Res}\left[\frac{e^{tz}}{z\sinh(az)};0\right] = \frac{1}{1!}\lim_{z\to 0}\frac{d}{dz}\left[\frac{(z-0)^2e^{tz}}{z\sinh(az)}\right] = \lim_{z\to 0}\frac{d}{dz}\left[\frac{ze^{tz}}{\sinh(az)}\right] \tag{12.9.15}$$

$$= \lim_{z\to 0}\left[\frac{e^{tz}}{\sinh(az)} + \frac{zte^{tz}}{\sinh(az)} - \frac{az\cosh(az)e^{tz}}{\sinh^2(az)}\right] = \frac{t}{a} \tag{12.9.16}$$

---

[33] Gradshteyn, I. S., and I. M. Ryzhik, 1965: *Table of Integrals, Series and Products*. Academic Press, Section 1.431, Formula 2.

after using $\sinh(az) = az + O(z^3)$. For the simple poles $z_n = \pm n\pi i/a$,

$$\operatorname{Res}\left[\frac{e^{tz}}{z\sinh(az)}; z_n\right] = \lim_{z\to z_n} \frac{(z-z_n)e^{tz}}{z\sinh(az)} = \lim_{z\to z_n} \frac{e^{tz}}{\sinh(az) + az\cosh(az)} \tag{12.9.17}$$

$$= \frac{\exp(\pm n\pi it/a)}{(-1)^n(\pm n\pi i)}, \tag{12.9.18}$$

because $\cosh(\pm n\pi i) = \cos(n\pi) = (-1)^n$. Thus, summing up all of the residues gives

$$f(t) = \frac{t}{a} + \sum_{n=1}^{\infty} \frac{(-1)^n \exp(n\pi it/a)}{n\pi i} - \sum_{n=1}^{\infty} \frac{(-1)^n \exp(-n\pi it/a)}{n\pi i} \tag{12.9.19}$$

$$= \frac{t}{a} + \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin(n\pi t/a). \tag{12.9.20}$$

□

In addition to computing the inverse of Laplace transforms, Bromwich's integral places certain restrictions on $F(s)$ in order that an inverse exists. If α denotes the minimum value that c may possess, the restrictions are threefold.[34] First, $F(z)$ must be analytic in the half-plane $x \ge \alpha$, where $z = x + iy$. Second, in the same half-plane it must behave as z^{-k}, where $k > 1$. Finally, $F(x)$ must be real when $x \ge \alpha$.

• Example 12.9.3

Is the function $\sin(s)/(s^2 + 4)$ a proper Laplace transform? Although the function satisfies the first and third criteria listed in the previous paragraph on the half-plane $x > 2$, the function becomes unbounded as $y \to \pm\infty$ for any fixed $x > 2$. Thus, $\sin(s)/(s^2 + 4)$ cannot be a Laplace transform. □

• Example 12.9.4

An additional benefit of understanding inversion by the residue method is the ability to qualitatively anticipate the inverse by knowing the location of the poles of $F(s)$. This intuition is important because many engineering analyses discuss stability and performance entirely in terms of the properties of the system's Laplace transform. In Figure 12.9.3 we have graphed the location of the poles of $F(s)$ and the corresponding $f(t)$. The student should go through the mental exercise of connecting the two pictures.

Problems

Use Bromwich's integral to invert the following Laplace transforms:

1. $F(s) = \dfrac{s+1}{(s+2)^2(s+3)}$

2. $F(s) = \dfrac{1}{s^2(s+a)^2}$

[34] For the proof, see Churchill, R. V., 1972: *Operational Mathematics*. McGraw-Hill, Section 67.

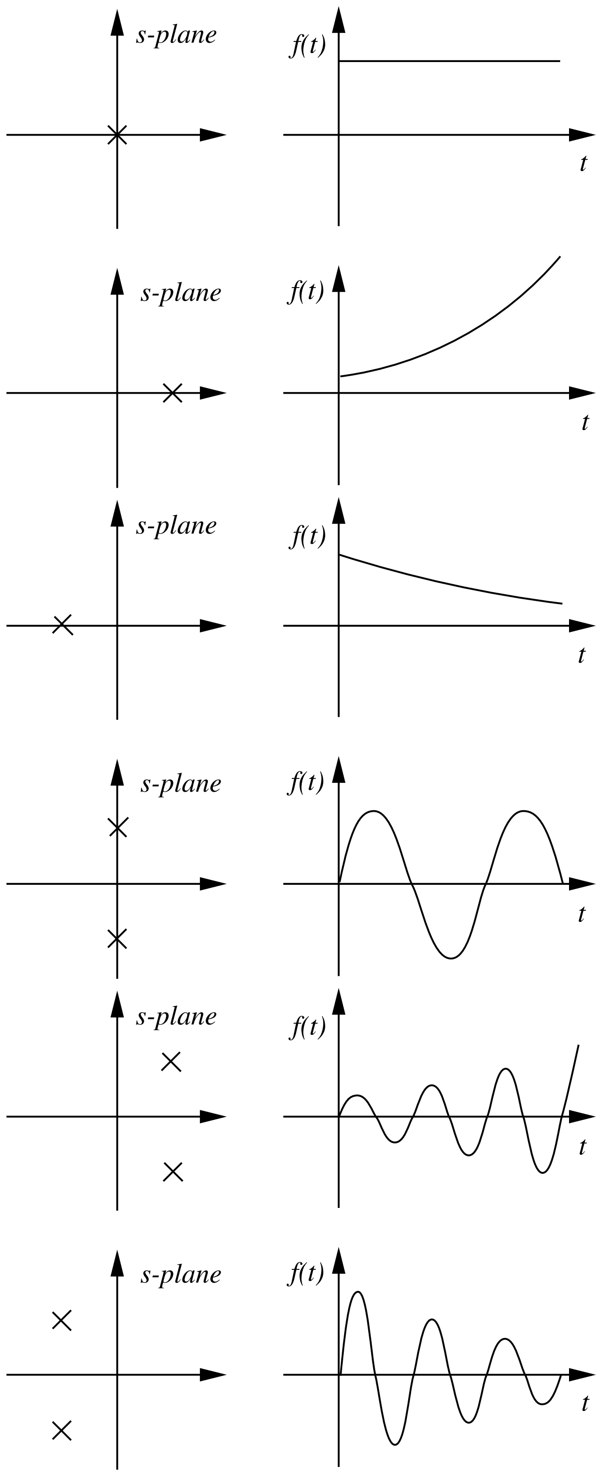


Figure 12.9.3: The correspondence between the location of the simple poles of the Laplace transform $F(s)$ and the behavior of $f(t)$.

3. $F(s) = \dfrac{1}{s(s-2)^3}$

4. $F(s) = \dfrac{1}{s(s+a)^2(s^2+b^2)}$

5. $F(s) = \dfrac{e^{-s}}{s^2(s+2)}$

6. $F(s) = \dfrac{1}{s(1+e^{-as})}$

7. $F(s) = \dfrac{1}{(s+b)\cosh(as)}$

8. $F(s) = \dfrac{1}{s(1-e^{-as})}$

9. Consider a function $f(t)$ that has the Laplace transform $F(z)$, which is analytic in the half-plane $\mathrm{Re}(z) > s_0$. Can we use this knowledge to find $g(t)$, whose Laplace transform $G(z)$ equals $F[\varphi(z)]$, where $\varphi(z)$ is also analytic for $\mathrm{Re}(z) > s_0$? The answer to this question leads to the Schouten[35]–Van der Pol[36] theorem.

Step 1: Show that the following relationships hold true:

$$G(z) = F[\varphi(z)] = \int_0^\infty f(\tau)e^{-\varphi(z)\tau}\, d\tau, \qquad \text{and} \qquad g(t) = \frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} F[\varphi(z)]e^{tz}\, dz.$$

Step 2: Using the results from Step 1, show that

$$g(t) = \int_0^\infty f(\tau)\left[\frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} e^{-\varphi(z)\tau}e^{tz}\, dz\right] d\tau.$$

This is the Schouten-Van der Pol theorem.

Step 3: If $G(z) = F(\sqrt{z}\,)$ show that

$$g(t) = \frac{1}{2\sqrt{\pi t^3}}\int_0^\infty \tau f(\tau)\exp\left(-\frac{\tau^2}{4t}\right) d\tau.$$

Hint: Do not evaluate the contour integral. Instead, ask yourself: What function of time has a Laplace transform that equals $e^{-\varphi(z)\tau}$, where τ is a parameter? Then use tables.

12.10 THE SOLUTION OF THE WAVE EQUATION

The solution of linear partial differential equations by Laplace transforms is the most commonly employed analytic technique after separation of variables. Because the transform consists solely of an integration with respect to time, the transform $U(x,s)$ of the solution of the wave equation $u(x,t)$ is

$$U(x,s) = \int_0^\infty u(x,t)e^{-st}\, dt, \qquad \textbf{(12.10.1)}$$

assuming that the wave equation only varies in a single spatial variable x and time t.

---

[35] Schouten, J. P., 1935: A new theorem in operational calculus together with an application of it. *Physica*, **2**, 75–80.

[36] Van der Pol, B., 1934: A theorem on electrical networks with applications to filters. *Physica*, **1**, 521–530.

Partial derivatives involving time have transforms similar to those that we encountered in the case of functions of a single variable. They include

$$\mathcal{L}[u_t(x,t)] = sU(x,s) - u(x,0), \tag{12.10.2}$$

and

$$\mathcal{L}[u_{tt}(x,t)] = s^2U(x,s) - su(x,0) - u_t(x,0). \tag{12.10.3}$$

These transforms introduce the initial conditions via $u(x,0)$ and $u_t(x,0)$. On the other hand, derivatives involving x become

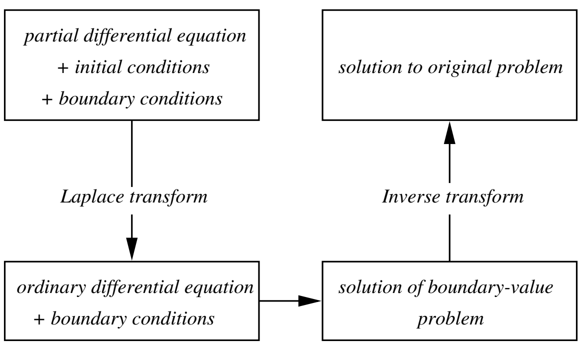
$$\mathcal{L}[u_x(x,t)] = \frac{d}{dx}\left\{\mathcal{L}[u(x,t)]\right\} = \frac{dU(x,s)}{dx}, \tag{12.10.4}$$

and

$$\mathcal{L}[u_{xx}(x,t)] = \frac{d^2}{dx^2}\left\{\mathcal{L}[u(x,t)]\right\} = \frac{d^2U(x,s)}{dx^2}. \tag{12.10.5}$$

Because the transformation eliminates the time variable, only $U(x,s)$ and its derivatives remain in the equation. Consequently, we transform the partial differential equation into a boundary-value problem involving an ordinary differential equation. Because this equation is often easier to solve than a partial differential equation, the use of Laplace transforms considerably simplifies the original problem. Of course, the Laplace transforms must exist for this technique to work.

The following schematic summarizes the Laplace transform method:



In the following examples, we illustrate transform methods by solving the classic equation of telegraphy as it applies to a uniform transmission line. The line has a resistance R, an inductance L, a capacitance C, and a leakage conductance G per unit length. We denote the current in the direction of positive x by I; V is the voltage drop across the transmission line at the point x. The dependent variables I and V are functions of both distance x along the line and time t.

To derive the differential equations that govern the current and voltage in the line, consider the points A at x and B at $x+\Delta x$ in Figure 12.10.1. The current and voltage at A are $I(x,t)$ and $V(x,t)$; at B, $I+\frac{\partial I}{\partial x}\Delta x$ and $V+\frac{\partial V}{\partial x}\Delta x$. Therefore, the voltage drop from A to B is $-\frac{\partial V}{\partial x}\Delta x$ and the current in the line is $I+\frac{\partial I}{\partial x}\Delta x$. Neglecting terms that are proportional to $(\Delta x)^2$,

$$\left(L\frac{\partial I}{\partial t} + RI\right)\Delta x = -\frac{\partial V}{\partial x}\Delta x. \tag{12.10.6}$$

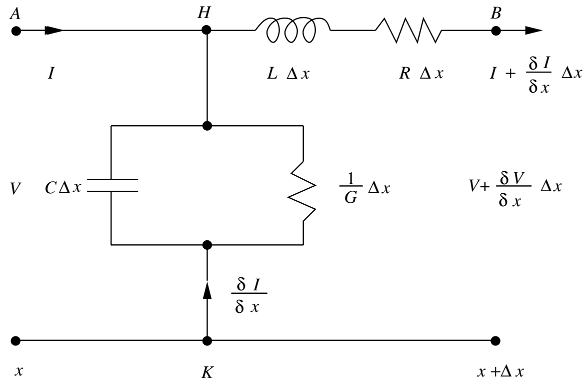


Figure 12.10.1: Schematic of a uniform transmission line.

The voltage drop over the parallel portion HK of the line is V while the current in this portion of the line is $-\frac{\partial I}{\partial x}\Delta x$. Thus,

$$\left(C\frac{\partial V}{\partial t} + GV\right)\Delta x = -\frac{\partial I}{\partial x}\Delta x. \qquad \textbf{(12.10.7)}$$

Therefore, the differential equations for I and V are

$$L\frac{\partial I}{\partial t} + RI = -\frac{\partial V}{\partial x}, \qquad \textbf{(12.10.8)}$$

and

$$C\frac{\partial V}{\partial t} + GV = -\frac{\partial I}{\partial x}. \qquad \textbf{(12.10.9)}$$

Turning to the initial conditions, we solve these simultaneous partial differential equations with the initial conditions

$$I(x,0) = I_0(x), \qquad \textbf{(12.10.10)}$$

and

$$V(x,0) = V_0(x) \qquad \textbf{(12.10.11)}$$

for $0 < t$. There are also boundary conditions at the ends of the line; we will introduce them for each specific problem. For example, if the line is short-circuited at $x = a$, $V = 0$ at $x = a$; if there is an open circuit at $x = a$, $I = 0$ at $x = a$.

To solve Equation 12.10.8 and Equation 12.10.9 by Laplace transforms, we take the Laplace transform of both sides of these equations, which yields

$$(Ls+R)\overline{I}(x,s) = -\frac{d\overline{V}(x,s)}{dx} + LI_0(x), \qquad \textbf{(12.10.12)}$$

and

$$(Cs+G)\overline{V}(x,s) = -\frac{d\overline{I}(x,s)}{dx} + CV_0(x). \qquad \textbf{(12.10.13)}$$

Eliminating $\overline{I}$ gives an ordinary differential equation in $\overline{V}$

$$\frac{d^2\overline{V}}{dx^2} - q^2\overline{V} = L\frac{dI_0(x)}{dx} - C(Ls+R)V_0(x), \qquad \textbf{(12.10.14)}$$

where $q^2 = (Ls + R)(Cs + G)$. After finding $\overline{V}$, we may compute $\overline{I}$ from

$$\overline{I} = -\frac{1}{Ls + R}\frac{d\overline{V}}{dx} + \frac{LI_0(x)}{Ls + R}. \quad \textbf{(12.10.15)}$$

At this point we treat several classic cases.

• Example 12.10.1: The semi-infinite transmission line

We consider the problem of a semi-infinite line $0 < x$ with no initial current and charge. The end $x = 0$ has a constant voltage E for $0 < t$.

In this case,

$$\frac{d^2\overline{V}}{dx^2} - q^2\overline{V} = 0, \qquad 0 < x. \quad \textbf{(12.10.16)}$$

The boundary conditions at the ends of the line are

$$V(0, t) = E, \qquad 0 < t, \quad \textbf{(12.10.17)}$$

and $V(x, t)$ is finite as $x \to \infty$. The transform of these boundary conditions is

$$\overline{V}(0, s) = E/s, \qquad \text{and} \qquad \lim_{x\to\infty} \overline{V}(x, s) \to 0. \quad \textbf{(12.10.18)}$$

The general solution of Equation 12.10.16 is

$$\overline{V}(x, s) = Ae^{-qx} + Be^{qx}. \quad \textbf{(12.10.19)}$$

The requirement that $\overline{V}$ remains finite as $x \to \infty$ forces $B = 0$. The boundary condition at $x = 0$ gives $A = E/s$. Thus,

$$\overline{V}(x, s) = \frac{E}{s}\exp\left[-\sqrt{(Ls + R)(Cs + G)}\,x\right]. \quad \textbf{(12.10.20)}$$

We discuss the general case later. However, for the so-called "lossless" line, where $R = G = 0$,

$$\overline{V}(x, s) = \frac{E}{s}\exp(-sx/c), \quad \textbf{(12.10.21)}$$

where $c = 1/\sqrt{LC}$. Consequently,

$$V(x, t) = EH\left(t - \frac{x}{c}\right), \quad \textbf{(12.10.22)}$$

where $H(t)$ is Heaviside's step function. The physical interpretation of this solution is as follows: $V(x, t)$ is zero up to the time x/c at which time a wave traveling with speed c from $x = 0$ would arrive at the point x. $V(x, t)$ has the constant value E afterwards.

For the so-called "distortionless" line,[37] $R/L = G/C = \rho$,

$$V(x, t) = Ee^{-\rho x/c}H\left(t - \frac{x}{c}\right). \quad \textbf{(12.10.23)}$$

---

[37] Prechtl and Schürhuber (Prechtl, A., and R. Schürhuber, 2000: Nonuniform distortionless transmission lines. *Electr. Eng. [Berlin]*, **82**, 127–134) generalized this problem to nonuniform transmission lines.

In this case, the disturbance not only propagates with velocity c but also attenuates as we move along the line.

Suppose now, that instead of applying a constant voltage E at $x = 0$, we apply a time-dependent voltage, $f(t)$. The only modification is that in place of Equation 12.10.20,

$$\overline{V}(x,s) = F(s)e^{-qx}. \tag{12.10.24}$$

In the case of the distortionless line, $q = (s+\rho)/c$, this becomes

$$\overline{V}(x,s) = F(s)e^{-(s+\rho)x/c} \tag{12.10.25}$$

and

$$V(x,t) = e^{-\rho x/c} f\left(t - \frac{x}{c}\right) H\left(t - \frac{x}{c}\right). \tag{12.10.26}$$

Thus, our solution shows that the voltage at x is zero up to the time x/c. Afterwards $V(x,t)$ follows the voltage at $x = 0$ with a time lag of x/c and decreases in magnitude by $e^{-\rho x/c}$.□

• Example 12.10.2: The finite transmission line

We now discuss the problem of a finite transmission line $0 < x < l$ with zero initial current and charge. We ground the end $x = 0$ and maintain the end $x = l$ at constant voltage E for $0 < t$.

The transformed partial differential equation becomes

$$\frac{d^2\overline{V}}{dx^2} - q^2\overline{V} = 0, \qquad 0 < x < l. \tag{12.10.27}$$

The boundary conditions are

$$V(0,t) = 0, \quad \text{and} \quad V(l,t) = E, \qquad 0 < t. \tag{12.10.28}$$

The Laplace transform of these boundary conditions is

$$\overline{V}(0,s) = 0, \quad \text{and} \quad \overline{V}(l,s) = E/s. \tag{12.10.29}$$

The solution of Equation 12.10.27 that satisfies the boundary conditions is

$$\overline{V}(x,s) = \frac{E\,\sinh(qx)}{s\,\sinh(ql)}. \tag{12.10.30}$$

Let us rewrite Equation 12.10.30 in a form involving negative exponentials and expand the denominator by the binomial theorem,

$$\overline{V}(x,s) = \frac{E}{s}e^{-q(l-x)}\frac{1-e^{-2qx}}{1-e^{-2ql}} \tag{12.10.31}$$

$$= \frac{E}{s}e^{-q(l-x)}(1-e^{-2qx})\left(1+e^{-2ql}+e^{-4ql}+\cdots\right) \tag{12.10.32}$$

$$= \frac{E}{s}\left[e^{-q(l-x)} - e^{-q(l+x)} + e^{-q(3l-x)} - e^{-q(3l+x)} + \cdots\right]. \tag{12.10.33}$$

In the special case of the lossless line where $q = s/c$,

$$\overline{V}(x,s) = \frac{E}{s}\left[e^{-s(l-x)/c} - e^{-s(l+x)/c} + e^{-s(3l-x)/c} - e^{-s(3l+x)/c} + \cdots\right], \tag{12.10.34}$$

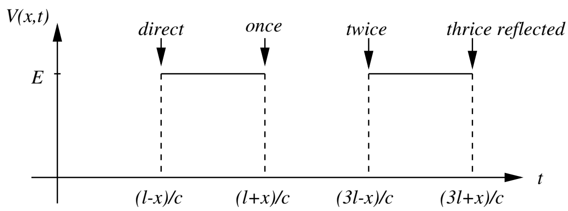


Figure 12.10.2: The voltage within a lossless, finite transmission line of length l as a function of time t.

or

$$V(x,t) = E\left[H\left(t - \frac{l-x}{c}\right) - H\left(t - \frac{l+x}{c}\right) + H\left(t - \frac{3l-x}{c}\right) - H\left(t - \frac{3l+x}{c}\right) + \cdots\right]. \tag{12.10.35}$$

We illustrate Equation 12.10.35 in Figure 12.10.2. The voltage at x is zero up to the time $(l-x)/c$, at which time a wave traveling directly from the end $x = l$ would reach the point x. The voltage then has the constant value E up to the time $(l+x)/c$, at which time a wave traveling from the end $x = l$ and reflected back from the end $x = 0$ would arrive. From this time up to the time of arrival of a twice-reflected wave, it has the value zero, and so on. □

• **Example 12.10.3: The semi-infinite transmission line reconsidered**

In the first example, we showed that the transform of the solution for the semi-infinite line is

$$\overline{V}(x,s) = \frac{E}{s}e^{-qx}, \tag{12.10.36}$$

where $q^2 = (Ls+R)(Cs+G)$. In the case of a lossless line $(R = G = 0)$, we found traveling wave solutions.

In this example, we shall examine the case of a submarine cable,[38] where $L = G = 0$. In this special case,

$$\overline{V}(x,s) = \frac{E}{s}e^{-x\sqrt{s/\kappa}}, \tag{12.10.37}$$

where $\kappa = 1/(RC)$. From a table of Laplace transforms,[39] we can immediately invert Equation 12.10.37 and find that

$$V(x,t) = E\ \text{erfc}\left(\frac{x}{2\sqrt{\kappa t}}\right), \tag{12.10.38}$$

where erfc is the complementary error function. Unlike the traveling wave solution, the voltage diffuses into the cable as time increases. We illustrate Equation 12.10.38 in Figure 12.10.3. □

---

[38] First solved by Thomson, W., 1855: On the theory of the electric telegraph. *Proc. R. Soc. London, Ser. A*, **7**, 382–399.

[39] See Churchill, R. V., 1972: *Operational Mathematics*. McGraw-Hill Book, Section 27.

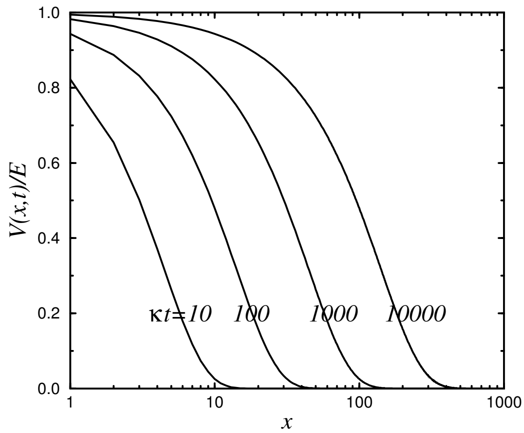


Figure 12.10.3: The voltage within a submarine cable as a function of distance for various values of κt.

• Example 12.10.4: A short-circuited, finite transmission line

Let us find the voltage of a lossless transmission line of length l that initially has the constant voltage E. At $t = 0$, we ground the line at $x = 0$ while we leave the end $x = l$ insulated.

The transformed partial differential equation now becomes

$$\frac{d^2\overline{V}}{dx^2} - \frac{s^2}{c^2}\overline{V} = -\frac{sE}{c^2}, \tag{12.10.39}$$

where $c = 1/\sqrt{LC}$. The boundary conditions are

$$\overline{V}(0, s) = 0, \tag{12.10.40}$$

and

$$\overline{I}(l, s) = -\frac{1}{Ls}\frac{d\overline{V}(l, s)}{dx} = 0 \tag{12.10.41}$$

from Equation 12.10.15.

The solution to this boundary-value problem is

$$\overline{V}(x, s) = \frac{E}{s} - \frac{E\cosh[s(l-x)/c]}{s\cosh(sl/c)}. \tag{12.10.42}$$

The first term on the right side of Equation 12.10.42 is easy to invert and the inversion equals E. The second term is much more difficult to handle. We will use Bromwich's integral.

In Section 12.9 we showed that

$$\mathcal{L}^{-1}\left\{\frac{\cosh[s(l-x)/c]}{s\cosh(sl/c)}\right\} = \frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i}\frac{\cosh[z(l-x)/c]e^{tz}}{z\cosh(zl/c)}\,dz. \tag{12.10.43}$$

To evaluate this integral we must first locate and then classify the singularities. Using the product formula for the hyperbolic cosine,

$$\frac{\cosh[z(l-x)/c]}{z\cosh(zl/c)} = \frac{[1 + \frac{4z^2(l-x)^2}{c^2\pi^2}][1 + \frac{4z^2(l-x)^2}{9c^2\pi^2}]\cdots}{z[1 + \frac{4z^2l^2}{c^2\pi^2}][1 + \frac{4z^2l^2}{9c^2\pi^2}]\cdots}. \tag{12.10.44}$$

This shows that we have an infinite number of simple poles located at $z = 0$, and $z_n = \pm(2n-1)\pi ci/(2l)$, where $n = 1, 2, 3, \ldots$. Therefore, Bromwich's contour can lie along, and just to the right of, the imaginary axis. By Jordan's lemma we close the contour with a semicircle of infinite radius in the left half of the complex plane. Computing the residues,

$$\operatorname{Res}\left\{\frac{\cosh[z(l-x)/c]e^{tz}}{z\,\cosh(zl/c)}; 0\right\} = \lim_{z\to 0}\frac{\cosh[z(l-x)/c]e^{tz}}{\cosh(zl/c)} = 1, \tag{12.10.45}$$

and

$$\operatorname{Res}\left\{\frac{\cosh[z(l-x)/c]e^{tz}}{z\,\cosh(zl/c)}; z_n\right\} = \lim_{z\to z_n}\frac{(z-z_n)\cosh[z(l-x)/c]e^{tz}}{z\,\cosh(zl/c)} \tag{12.10.46}$$

$$= \frac{\cosh[(2n-1)\pi(l-x)i/(2l)]\exp[\pm(2n-1)\pi cti/(2l)]}{[(2n-1)\pi i/2]\,\sinh[(2n-1)\pi i/2]} \tag{12.10.47}$$

$$= \frac{2(-1)^n}{(2n-1)\pi}\cos\left[\frac{(2n-1)\pi(l-x)}{2l}\right]\exp\left[\pm\frac{(2n-1)\pi cti}{2l}\right]. \tag{12.10.48}$$

Summing the residues and using the relationship that $\cos(t) = (e^{ti} + e^{-ti})/2$,

$$V(x,t) = E - E\left\{1 - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{2n-1}\cos\left[\frac{(2n-1)\pi(l-x)}{2l}\right]\cos\left[\frac{(2n-1)\pi ct}{2l}\right]\right\} \tag{12.10.49}$$

$$= \frac{4E}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{2n-1}\cos\left[\frac{(2n-1)\pi(l-x)}{2l}\right]\cos\left[\frac{(2n-1)\pi ct}{2l}\right]. \tag{12.10.50}$$

An alternative to contour integration is to rewrite Equation 12.10.42 as

$$\overline{V}(x,s) = \frac{E}{s}\left\{1 - \frac{e^{-sx/c}\left[1+e^{-2s(l-x)/c}\right]}{1+e^{-2sl/c}}\right] \tag{12.10.51}$$

$$= \frac{E}{s}\left[1 - e^{-sx/c} - e^{-s(2l-x)/c} + e^{-s(2l+x)/c} + \cdots\right] \tag{12.10.52}$$

so that

$$V(x,t) = E\left[1 - H\left(t-\frac{x}{c}\right) - H\left(t-\frac{2l-x}{c}\right) + H\left(t-\frac{2l+x}{c}\right) + \cdots\right]. \tag{12.10.53}$$

□

• Example 12.10.5: The general solution of the equation of telegraphy

In this example we solve the equation of telegraphy without any restrictions on R, C, G, or L. We begin by eliminating the dependent variable $I(x,t)$ from the set of equations, Equation 12.10.8 and Equation 12.10.9. This yields

$$CL\frac{\partial^2 V}{\partial t^2} + (GL+RC)\frac{\partial V}{\partial t} + RGV = \frac{\partial^2 V}{\partial x^2}. \tag{12.10.54}$$

We next take the Laplace transform of Equation 12.10.54 assuming that $V(x,0) = f(x)$, and $V_t(x,0) = g(x)$. The transformed version of Equation 12.10.54 is

$$\frac{d^2\overline{V}}{dx^2} - [CLs^2 + (GL+RC)s + RG]\overline{V} = -CLg(x) - (CLs + GL + RC)f(x), \quad \mathbf{(12.10.55)}$$

or

$$\frac{d^2\overline{V}}{dx^2} - \frac{(s+\rho)^2 - \sigma^2}{c^2}\overline{V} = -\frac{g(x)}{c^2} - \left(\frac{s}{c^2} + \frac{2\rho}{c^2}\right) f(x), \quad \mathbf{(12.10.56)}$$

where $c^2 = 1/LC$, $\rho = c^2(RC+GL)/2$, and $\sigma = c^2(RC-GL)/2$.

We solve Equation 12.10.56 by Fourier transforms (see Section 11.6) with the requirement that the solution dies away as $|x| \to \infty$. The most convenient way of expressing this solution is the convolution product (see Section 11.5)

$$\overline{V}(x,s) = \left[\frac{g(x)}{c} + \left(\frac{s}{c} + \frac{2\rho}{c}\right) f(x)\right] * \frac{\exp[-|x|\sqrt{(s+\rho)^2 - \sigma^2}/c]}{2\sqrt{(s+\rho)^2-\sigma^2}}. \quad \mathbf{(12.10.57)}$$

From a table of Laplace transforms,

$$\mathcal{L}^{-1}\left[\frac{\exp\left(-b\sqrt{s^2-a^2}\,\right)}{\sqrt{s^2-a^2}}\right] = I_0\left(a\sqrt{t^2-b^2}\,\right) H(t-b), \quad \mathbf{(12.10.58)}$$

where $b > 0$ and $I_0(\ )$ is the zeroth order modified Bessel function of the first kind. Therefore, by the first shifting theorem,

$$\mathcal{L}^{-1}\left\{\frac{\exp\left[-|x|\sqrt{(s+\rho)^2-\sigma^2}/c\right]}{\sqrt{(s+\rho)^2-\sigma^2}}\right\} = e^{-\rho t} I_0\left[\sigma\sqrt{t^2-(x/c)^2}\right] H\left(t - \frac{|x|}{c}\right). \quad \mathbf{(12.10.59)}$$

Using Equation 12.10.59 to invert Equation 12.10.57, we have that

$$\begin{aligned} V(x,t) = {} & \tfrac{1}{2c} e^{-\rho t} g(x) * I_0\left[\sigma\sqrt{t^2-(x/c)^2}\right] H(t-|x|/c) \\ & + \tfrac{1}{2c} e^{-\rho t} f(x) * \frac{\partial}{\partial t}\left\{I_0[\sigma\sqrt{t^2-(x/c)^2}]\right\} H(t-|x|/c) \\ & + \tfrac{\rho}{c} e^{-\rho t} f(x) * I_0\left[\sigma\sqrt{t^2-(x/c)^2}\right] H(t-|x|/c) \\ & + \tfrac{1}{2} e^{-\rho t}[f(x+ct) + f(x-ct)]. \end{aligned} \quad \mathbf{(12.10.60)}$$

The last term in Equation 12.10.60 arises from noting that $sF(s) = \mathcal{L}[f(t)] + f(0)$. If we explicitly write out the convolution, the final form of the solution is

$$\begin{aligned} V(x,t) = {} & \tfrac{1}{2} e^{-\rho t}[f(x+ct) + f(x-ct)] \\ & + \tfrac{1}{2c} e^{-\rho t} \int_{x-ct}^{x+ct} [g(\eta) + 2\rho f(\eta)] I_0\left[\sigma\sqrt{c^2t^2-(x-\eta)^2}\Big/ c\right] d\eta \\ & + \tfrac{1}{2c} e^{-\rho t} \int_{x-ct}^{x+ct} f(\eta) \frac{\partial}{\partial t}\left\{I_0\left[\sigma\sqrt{c^2t^2-(x-\eta)^2}\Big/ c\right]\right\} d\eta. \end{aligned} \quad \mathbf{(12.10.61)}$$

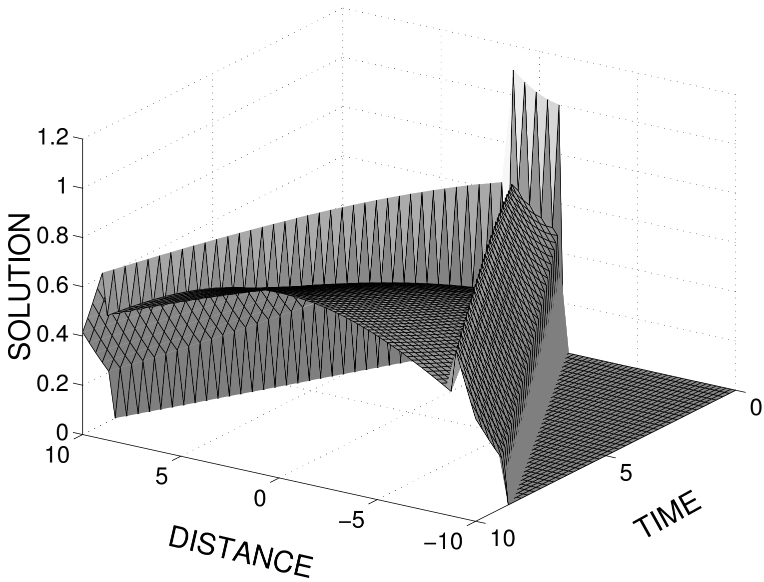


Figure 12.10.4: The evolution of the voltage with time given by the general equation of telegraphy for initial conditions and parameters stated in the text.

There is a straightforward physical interpretation of the first line of Equation 12.10.61. It represents damped progressive waves; one is propagating to the right and the other to the left. In addition to these progressive waves, there is a contribution from the integrals, even after the waves pass. These integrals include all of the points where $f(x)$ and $g(x)$ are nonzero within a distance ct from the point in question. This effect persists through all time, although dying away, and constitutes a residue or tail. Figure 12.10.4 illustrates this for $\rho = 0.1$, $\sigma = 0.2$, and $c = 1$. This figure was obtained using the MATLAB script:

```
% initialize parameters in calculation
clear; dx = 0.1; dt = 0.5; rho_over_c = 0.1; sigma_over_c = 0.2;
X=[-10:dx:10]; T = [0:dt:10]; % compute locations of x and t
for j=1:length(T); t = T(j);
for i=1:length(X); x = X(i);
  XX(i,j) = x; TT(i,j) = t; deta_i = 0.05 % set up grid
% compute characteristics x+ct and x-ct
  characteristic_1 = x - t; characteristic_2 = x + t;
% compute first term in Equation 12.10.61
  F = inline('stepfun(x,-1.0001)-stepfun(x,1.0001)');
  u(i,j ) = F(characteristic_1) + F(characteristic_2);
% find the upper and lower limits of the integration
  upper = characteristic_2; lower = characteristic_1;
  if t > 0 & upper > -1 & lower < 1
  if upper > 1 upper = 1; end
  if lower < -1 lower = -1; end
% set up parameters needed for integration
  interval = upper-lower;
  NN = interval / deta_i;
  if mod(NN,2) > 0 NN = NN + 1; end;
```

```
  deta = interval / NN;
% compute integrals in Equation 12.10.61 by Simpson's rule
% sum1 deals with the first integral while sum2 is the second
  sum1 = 0; sum2 = 0; eta = lower;
  for k = 0:2:NN-2
    arg = sigma_over_c * sqrt(t*t-(x-eta)*(x-eta));
    sum1 = sum1 + besseli(0,arg);
    if (arg == 0)
      sum2 = sum2 + 0.5 * sigma_over_c * t;
    else
      sum2 = sum2 + t * besseli(1,arg) / arg; end
    eta = eta + deta;
    arg = sigma_over_c * sqrt(t*t-(x-eta)*(x-eta));
    sum1 = sum1 + 4*besseli(0,arg);
    if (arg == 0)
      sum2 = sum2 + 4 * 0.5 * sigma_over_c * t;
    else
      sum2 = sum2 + 4 * t * besseli(1,arg) / arg; end
    eta = eta + deta;
    arg = sigma_over_c * sqrt(t*t-(x-eta)*(x-eta));
    sum1 = sum1 + besseli(0,arg);
    if (arg == 0)
      sum2 = sum2 + 0.5 * sigma_over_c * t;
    else
      sum2 = sum2 + t * besseli(1,arg) / arg; end
    end
  u(i,j) = u(i,j) + 2 * rho_over_c * deta * sum1 / 3 ...
         + sigma_over_c * deta * sum2 / 3;
  end
% multiply final answer by damping coefficient
  u(i,j) = 0.5 * exp(-rho_over_c * t) * u(i,j);
end;end;
% plot results
mesh(XX,TT,real(u)); colormap spring;
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

We evaluated the integrals by Simpson's rule for the initial conditions $f(x) = H(x+1) - H(x-1)$, and $g(x) = 0$. If there was no loss, then two pulses would propagate to the left and right. However, with resistance and leakage the waves leave a residue after their leading edge has passed.

Problems

1. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

with the boundary conditions $u(0,t) = u(1,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = 0$, $u_t(0,t) = 1$, $0 < x < 1$.

2. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

with the boundary conditions $u(0,t) = u_x(1,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = 0$, $u_t(0,t) = x$, $0 < x < 1$.

3. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

with the boundary conditions $u(0,t) = u(1,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = \sin(\pi x)$, $u_t(x,0) = -\sin(\pi x)$, $0 < x < 1$.

4. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < a, \quad 0 < t,$$

with the boundary conditions $u(0,t) = \sin(\omega t)$, $u(a,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = u_t(x,0) = 0$, $0 < x < a$. Assume that $\omega a/c$ is *not* an integer multiple of π. Why?

5. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

with the boundary conditions $u_x(0,t) = -f(t)$, $u_x(L,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = u_t(x,0) = 0$, $0 < x < L$. Hint: Invert the Laplace transform following the procedure used in Example 12.10.2.

6. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} - q'(t), \qquad a < x < b, \quad 0 < t,$$

with the boundary conditions $u(a,t) = 0$, $u_x(b,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = 0$, $u_t(x,0) = -q(0)$, $a < x < b$. Hint: To find $U(x,s)$, express both $U(x,s)$ and the right side of the ordinary differential equation governing $U(x,s)$ in an eigenfunction expansion using $\sin\{(2n+1)\pi(x-a)/[2(b-a)]\}$. These eigenfunctions satisfy the boundary conditions.

7. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = te^{-x}, \qquad 0 < x < \infty, \quad 0 < t,$$

with the boundary conditions

$$u(0,t) = 1 - e^{-t}, \quad \lim_{x\to\infty} |u(x,t)| \sim x^n, \ n \text{ finite}, \qquad 0 < t,$$

and the initial conditions $u(x,0) = 0$, $u_t(x,0) = x$, $0 < x < \infty$.

8. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = xe^{-t}, \qquad 0 < x < \infty, \quad 0 < t,$$

with the boundary conditions

$$u(0,t) = \cos(t), \qquad \lim_{x\to\infty} |u(x,t)| \sim x^n, \ n \text{ finite}, \qquad 0 < t,$$

and the initial conditions $u(x,0) = 1$, $u_t(x,0) = 0$, $0 < x < \infty$.

9. Use transform methods to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

with the boundary conditions

$$u(0,t) = 0, \qquad \frac{\partial^2 u(L,t)}{\partial t^2} + \frac{k}{m}\frac{\partial u(L,t)}{\partial x} = g, \qquad 0 < t,$$

and the initial conditions $u(x,0) = u_t(x,0) = 0$, $0 < x < L$, where k, m, and g are constants.

10. Use transform methods[40] to solve the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial x}\right), \qquad 0 < x < 1, \quad 0 < t,$$

with the boundary conditions

$$\lim_{x\to 0} |u(x,t)| < \infty, \quad u(1,t) = A\sin(\omega t), \qquad 0 < t,$$

and the initial conditions $u(x,0) = u_t(x,0) = 0$, $0 < x < 1$. Assume that $2\omega \neq c\beta_n$, where $J_0(\beta_n) = 0$. Hint: The ordinary differential equation

$$\frac{d}{dx}\left(x\frac{dU}{dx}\right) - \frac{s^2}{c^2}U = 0$$

has the solution

$$U(x,s) = c_1 I_0\left(\frac{s}{c}\sqrt{x}\right) + c_2 K_0\left(\frac{s}{c}\sqrt{x}\right),$$

where $I_0(x)$ and $K_0(x)$ are modified Bessel functions of the first and second kind, respectively. Note that $J_n(iz) = i^n I_n(z)$ and $I_n(iz) = i^n J_n(z)$ for complex z.

11. A lossless transmission line of length ℓ has a constant voltage E applied to the end $x = 0$ while we insulate the other end $[V_x(\ell,t) = 0]$. Find the voltage at any point on the line if the initial current and charge are zero.

[40] Suggested by a problem solved by Brown, J., 1975: Stresses in towed cables during re-entry. *J. Spacecr. Rockets*, **12**, 524–527.

12. Solve the equation of telegraphy without leakage

$$\frac{\partial^2 u}{\partial x^2} = CR\frac{\partial u}{\partial t} + CL\frac{\partial^2 u}{\partial t^2}, \qquad 0 < x < \ell, \quad 0 < t,$$

subject to the boundary conditions $u(0,t) = 0$, $u(\ell,t) = E$, $0 < t$, and the initial conditions $u(x,0) = u_t(x,0) = 0$, $0 < x < \ell$. Assume that $4\pi^2 L/CR^2\ell^2 > 1$. Why?

13. The pressure and velocity oscillations from water hammer in a pipe without friction[41] are given by the equations

$$\frac{\partial p}{\partial t} = -\rho c^2 \frac{\partial u}{\partial x}, \qquad \text{and} \qquad \frac{\partial u}{\partial t} = -\frac{1}{\rho}\frac{\partial p}{\partial x},$$

where $p(x,t)$ denotes the pressure perturbation, $u(x,t)$ is the velocity perturbation, c is the speed of sound in water, and ρ is the density of water. These two first-order partial differential equations can be combined to yield

$$\frac{\partial^2 p}{\partial t^2} = c^2 \frac{\partial^2 p}{\partial x^2}.$$

Find the solution to this partial differential equation if $p(0,t) = p_0$, and $u(L,t) = 0$, and the initial conditions are $p(x,0) = p_0$, $p_t(x,0) = 0$, and $u(x,0) = u_0$.

14. Use Laplace transforms to solve the wave equation[42]

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} - \frac{2u}{r^2} \right), \qquad a < r < \infty, \quad 0 < t,$$

subject to the boundary conditions that

$$u(a,t) = A\left(1 - e^{-ct/a}\right) H(t), \quad \lim_{r\to\infty} u(r,t) \to 0, \qquad 0 < t,$$

and the initial conditions that $u(r,0) = u_t(r,0) = 0$, $a < r < \infty$. Hint: The homogeneous solution to the ordinary differential equation

$$\frac{d^2 y}{dr^2} + \frac{2}{r}\frac{dy}{dr} - \frac{2y}{r^2} - b^2 y = 0$$

is

$$y(r) = C_1 \left[\frac{\cosh(br)}{br} - \frac{\sinh(br)}{b^2 r^2} \right] + C_2 \left(\frac{1}{br} + \frac{1}{b^2 r^2} \right) e^{-br}.$$

15. Use Laplace transforms to solve the wave equation[43]

$$\frac{\partial^2 (ru)}{\partial t^2} = c^2 \frac{\partial^2 (ru)}{\partial r^2}, \qquad a < r < \infty, \quad 0 < t,$$

[41] See Rich, G. R., 1945: Water-hammer analysis by the Laplace-Mellin transformation. *Trans. ASME*, **67**, 361–376.

[42] Wolf, J. P., and G. R. Darbre, 1986: Time-domain boundary element method in visco-elasticity with application to a spherical cavity. *Soil Dynam. Earthq. Eng.*, **5**, 138–148.

[43] Originally solved using Fourier transforms by Sharpe, J. A., 1942: The production of elastic waves by explosion pressures. I. Theory and empirical field observations. *Geophysics*, **7**, 144–154.

subject to the boundary conditions that

$$-\rho c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{3r}\frac{\partial u}{\partial r} \right)\bigg|_{r=a} = p_0 e^{-\alpha t} H(t), \quad \lim_{r\to\infty} u(r,t) \to 0, \qquad 0 < t,$$

where $\alpha > 0$, and the initial conditions that $u(r,0) = u_t(r,0) = 0$, $a < r < \infty$.

16. Consider a vertical rod or column of length L that is supported at both ends. The elastic waves that arise when the support at the bottom is suddenly removed are governed by the wave equation[44]

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + g, \qquad 0 < x < L, \quad 0 < t,$$

where g denotes the gravitational acceleration, $c^2 = E/\rho$, E is Young's modulus, and ρ is the mass density. Find the wave solution if the boundary conditions are $u_x(0,t) = u_x(L,t) = 0$, $0 < t$, and the initial conditions are

$$u(x,0) = -\frac{gx^2}{2c^2}, \quad \frac{\partial u(x,0)}{\partial t} = 0, \qquad 0 < x < L.$$

17. Use Laplace transforms to solve the hyperbolic equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} + 1 = 0, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions that $u_x(0,t) = 0$, $u_x(1,t) = 1$, $0 < t$, and the initial conditions that $u(x,0) = u_t(x,0) = 0$, $0 < x < 1$.

18. Solve the telegraph-like equation[45]

$$\frac{\partial^2 u}{\partial t^2} + k\frac{\partial u}{\partial t} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \alpha \frac{\partial u}{\partial x} \right), \qquad 0 < x < \infty, \quad 0 \le t$$

subject to the boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = -u_0 \delta(t), \quad \lim_{x\to\infty} u(x,t) \to 0, \qquad 0 \le t,$$

and the initial conditions $u(x,0) = u_0$, $u_t(x,0) = 0$, $0 < x < \infty$, with $\alpha c > k$.

Step 1: Take the Laplace transform of the partial differential equation and boundary conditions and show that

$$\frac{d^2 U(x,s)}{dx^2} + \alpha \frac{dU(x,s)}{dx} - \left(\frac{s^2 + ks}{c^2} \right) U(x,s) = -\left(\frac{s+k}{c^2} \right) u_0,$$

---

[44] See Hall, L. H., 1953: Longitudinal vibrations of a vertical column by the method of Laplace transform. *Am. J. Phys.*, **21**, 287–292.

[45] See Abbott, M. R., 1959: The downstream effect of closing a barrier across an estuary with particular reference to the Thames. *Proc. R. Soc. London, Ser. A*, **251**, 426–439.

with $U'(0,s) = -u_0$, and $\lim_{x\to\infty} U(x,s) \to 0$.

Step 2: Show that the solution to the previous step is

$$U(x,s) = \frac{u_0}{s} + u_0 e^{-\alpha x/2}\,\frac{\exp\left[-x\sqrt{\left(s+\frac{k}{2}\right)^2 + a^2/c}\right]}{\frac{\alpha}{2} + \sqrt{(s+\frac{k}{2})^2 + a^2/c}},$$

where $4a^2 = \alpha^2 c^2 - k^2 > 0$.

Step 3: Using the first and second shifting theorems and the property that

$$F\left(\sqrt{s^2+a^2}\right) = \mathcal{L}\left[f(t) - a\int_0^t \frac{J_1\left(a\sqrt{t^2-\tau^2}\,\right)}{\sqrt{t^2-\tau^2}}\tau f(\tau)\,d\tau\right],$$

show that

$$u(x,t) = u_0 + u_0 c e^{-kt/2} H(t - x/c)\left[e^{-\alpha ct/2} - a\int_{x/c}^t \frac{J_1\left(a\sqrt{t^2-\tau^2}\,\right)}{\sqrt{t^2-\tau^2}}\tau e^{-\alpha c\tau/2}d\tau\right].$$

19. As an electric locomotive travels down a track at the speed V, the pantograph (the metallic framework that connects the overhead power lines to the locomotive) pushes up the line with a force P. Let us find the behavior[46] of the overhead wire as a pantograph passes between two supports of the electrical cable that are located a distance L apart. We model this system as a vibrating string with a point load:

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2} + \frac{P}{\rho V}\,\delta\Big(t - \frac{x}{V}\Big), \qquad 0 < x < L, \quad 0 < t.$$

Let us assume that the wire is initially at rest $[u(x,0) = u_t(x,0) = 0$ for $0 < x < L]$ and fixed at both ends $[u(0,t) = u(L,t) = 0$ for $0 < t]$.

Step 1: Take the Laplace transform of the partial differential equation and show that

$$s^2 U(x,s) = c^2\frac{d^2U(x,s)}{dx^2} + \frac{P}{\rho V}e^{-xs/V}.$$

Step 2: Solve the ordinary differential equation in Step 1 as a Fourier half-range sine series

$$U(x,s) = \sum_{n=1}^\infty B_n(s)\sin\Big(\frac{n\pi x}{L}\Big),$$

where

$$B_n(s) = \frac{2P\beta_n}{\rho L(\beta_n^2 - \alpha_n^2)}\left[\frac{1}{s^2+\alpha_n^2} - \frac{1}{s^2+\beta_n^2}\right]\left[1 - (-1)^n e^{-Ls/V}\right],$$

[46] See Oda, O., and Y. Ooura, 1976: Vibrations of catenary overhead wire. *Q. Rep., (Tokyo) Railway Tech. Res. Inst.*, **17**, 134–135.

$\alpha_n = n\pi c/L$ and $\beta_n = n\pi V/L$. This solution satisfies the boundary conditions.

Step 3: By inverting the solution in Step 2, show that

$$u(x,t) = \frac{2P}{\rho L}\sum_{n=1}^{\infty}\left[\frac{\sin(\beta_n t)}{\alpha_n^2-\beta_n^2} - \frac{V}{c}\,\frac{\sin(\alpha_n t)}{\alpha_n^2-\beta_n^2}\right]\sin\left(\frac{n\pi x}{L}\right)$$
$$- \frac{2P}{\rho L}H\left(t-\frac{L}{V}\right)\sum_{n=1}^{\infty}(-1)^n\sin\left(\frac{n\pi x}{L}\right)\left\{\frac{\sin[\beta_n(t-L/V)]}{\alpha_n^2-\beta_n^2} - \frac{V}{c}\,\frac{\sin[\alpha_n(t-L/V)]}{\alpha_n^2-\beta_n^2}\right\}$$

or

$$u(x,t) = \frac{2P}{\rho L}\sum_{n=1}^{\infty}\left[\frac{\sin(\beta_n t)}{\alpha_n^2-\beta_n^2} - \frac{V}{c}\,\frac{\sin(\alpha_n t)}{\alpha_n^2-\beta_n^2}\right]\sin\left(\frac{n\pi x}{L}\right)$$
$$- \frac{2P}{\rho L}H\left(t-\frac{L}{V}\right)\sum_{n=1}^{\infty}\sin\left(\frac{n\pi x}{L}\right)\left\{\frac{\sin(\beta_n t)}{\alpha_n^2-\beta_n^2} - \frac{V}{c}(-1)^n\frac{\sin[\alpha_n(t-L/V)]}{\alpha_n^2-\beta_n^2}\right\}.$$

The first term in both summations represents the static uplift on the line; this term disappears after the pantograph passes. The second term in both summations represents the vibrations excited by the traveling force. Even after the pantograph passes, they continue to exist.

20. Solve the wave equation

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial r^2} - \frac{1}{r}\frac{\partial u}{\partial r} + \frac{u}{r^2} = \frac{\delta(r-\alpha)}{\alpha^2}, \qquad 0 \le r < a, \quad 0 < t,$$

where $0 < \alpha < a$, subject to the boundary conditions

$$\lim_{r\to 0}|u(r,t)| < \infty, \qquad \frac{\partial u(a,t)}{\partial r} + \frac{h}{a}u(a,t) = 0, \qquad 0 < t,$$

and the initial conditions $u(r,0) = u_t(r,0) = 0$, $0 \le r < a$.

Step 1: Take the Laplace transform of the partial differential equation and show that

$$\frac{d^2U(r,s)}{dr^2} + \frac{1}{r}\frac{dU(r,s)}{dr} - \left(\frac{s^2}{c^2}+\frac{1}{r^2}\right)U(r,s) = -\frac{\delta(r-\alpha)}{s\alpha^2}, \quad 0 \le r < a,$$

with

$$\lim_{r\to 0}|U(r,s)| < \infty, \qquad \frac{dU(a,s)}{dr} + \frac{h}{a}U(a,s) = 0.$$

Step 2: Show that the Dirac delta function can be reexpressed as the Fourier-Bessel series

$$\delta(r-\alpha) = \frac{2\alpha}{a^2}\sum_{n=1}^{\infty}\frac{\beta_n^2\,J_1(\beta_n\alpha/a)}{(\beta_n^2+h^2-1)J_1^2(\beta_n)}J_1(\beta_n r/a), \qquad 0 \le r < a,$$

where β_n is the nth root of $\beta J_1'(\beta) + h\,J_1(\beta) = \beta J_0(\beta) + (h-1)J_1(\beta) = 0$ and $J_0(\ ), J_1(\ )$ are the zeroth and first-order Bessel functions of the first kind, respectively.

Step 3: Show that the solution to the ordinary differential equation in Step 1 is

$$U(r,s) = \frac{2}{\alpha}\sum_{n=1}^{\infty}\frac{J_1(\beta_n\alpha/a)J_1(\beta_n r/a)}{(\beta_n^2+h^2-1)\,J_1^2(\beta_n)}\left[\frac{1}{s}-\frac{s}{s^2+c^2\beta_n^2/a^2}\right].$$

Note that this solution satisfies the boundary conditions.

Step 4: Taking the inverse of the Laplace transform in Step 3, show that the solution to the partial differential equation is

$$u(r,t) = \frac{2}{\alpha}\sum_{n=1}^{\infty}\frac{J_1(\beta_n\alpha/a)J_1(\beta_n r/a)}{(\beta_n^2+h^2-1)\,J_1^2(\beta_n)}\left[1-\cos\left(\frac{c\beta_n t}{a}\right)\right].$$

21. Solve the hyperbolic equation

$$\frac{\partial^2 u}{\partial x\partial t}+u=0, \qquad 0<x,t,$$

subject to the boundary conditions $u(0,t)=e^{-t}$, $\lim_{x\to\infty}u(x,t)\to 0$, $0<t$, and $u(x,0)=1$, $\lim_{t\to\infty}|u(x,t)|<Me^{kt}$, $0<k,M,x,t$.

Step 1: Take the Laplace transform of the partial differential equation and show that

$$s\frac{dU(x,s)}{dx}+U=0, \qquad U(0,s)=\frac{1}{s+1}, \quad \lim_{x\to\infty}U(x,s)\to 0.$$

Step 2: Show that

$$U(x,s)=\frac{e^{-x/s}}{s+1}=\frac{e^{-x/s}}{s}-\frac{e^{-x/s}}{s(s+1)}.$$

Step 3: Using tables and the convolution theorem, show that the solution is

$$u(x,t)=J_0(2\sqrt{xt}\,)-e^{-t}\int_0^t e^{\tau}J_0(2\sqrt{x\tau}\,)\,d\tau,$$

where $J_0(\cdot)$ is the Bessel function of the first kind and order zero.

22. Solve the hyperbolic equation

$$\frac{\partial^2 u}{\partial x\partial t}+a\frac{\partial u}{\partial t}+b\frac{\partial u}{\partial x}=0, \qquad 0<a,b,x,t,$$

subject to the boundary conditions $u(0,t)=e^{ct}$, $\lim_{x\to\infty}u(x,t)\to 0$, $0<t$, and the initial conditions $u(x,0)=1$, $\lim_{t\to\infty}|u(x,t)|<Me^{kt}$, $0<k,M,t,x$.

Step 1: Take the Laplace transform of the partial differential equation and show that

$$(s+b)\frac{dU(x,s)}{dx}+asU=a, \qquad U(0,s)=\frac{1}{s-c}, \quad \lim_{x\to\infty}U(x,s)\to 0.$$

Step 2: Show that

$$U(x,s) = \frac{1}{s} + \frac{c\,e^{-ax}}{s(s-c)} \exp\left(\frac{bx}{s+b}\right).$$

Step 3: Using tables, the first shifting theorem, and the convolution theorem, show that the solution is

$$u(x,t) = 1 + c\,e^{ct-ax} \int_0^t e^{-(b+c)\tau} I_0\left(2\sqrt{bx\tau}\right)\, d\tau,$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind and order zero.

12.11 THE SOLUTION OF THE HEAT EQUATION

In the previous section we showed that we can solve the wave equation by the method of Laplace transforms. This is also true for the heat equation. Once again, we take the Laplace transform with respect to time. From the definition of Laplace transforms,

$$\mathcal{L}[u(x,t)] = U(x,s), \tag{12.11.1}$$

$$\mathcal{L}[u_t(x,t)] = sU(x,s) - u(x,0), \tag{12.11.2}$$

and

$$\mathcal{L}[u_{xx}(x,t)] = \frac{d^2U(x,s)}{dx^2}. \tag{12.11.3}$$

We next solve the resulting ordinary differential equation, known as the *auxiliary equation*, along with the corresponding Laplace transformed boundary conditions. The initial condition gives us the value of $u(x,0)$. The final step is the inversion of the Laplace transform $U(x,s)$. We typically use the inversion integral.

• Example 12.11.1

To illustrate these concepts, we solve a heat conduction problem[47] in a plane slab of thickness $2L$. Initially the slab has a constant temperature of unity. For $0 < t$, we allow both faces of the slab to radiatively cool in a medium that has a temperature of zero.

If $u(x,t)$ denotes the temperature, a^2 is the thermal diffusivity, h is the relative emissivity, t is the time, and x is the distance perpendicular to the face of the slab and measured from the middle of the slab, then the governing equation is

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad -L < x < L, \quad 0 < t, \tag{12.11.4}$$

with the initial condition

$$u(x,0) = 1, \qquad -L < x < L, \tag{12.11.5}$$

and boundary conditions

$$\frac{\partial u(L,t)}{\partial x} + hu(L,t) = 0, \quad \frac{\partial u(-L,t)}{\partial x} + hu(-L,t) = 0, \quad 0 < t. \tag{12.11.6}$$

---

[47] Goldstein, S., 1932: The application of Heaviside's operational method to the solution of a problem in heat conduction. *Z. Angew. Math. Mech.*, **12**, 234–243.

Taking the Laplace transform of Equation 12.11.4 and substituting the initial condition,

$$a^2 \frac{d^2U(x,s)}{dx^2} - sU(x,s) = -1. \tag{12.11.7}$$

If we write $s = a^2q^2$, Equation 12.11.7 becomes

$$\frac{d^2U(x,s)}{dx^2} - q^2U(x,s) = -\frac{1}{a^2}. \tag{12.11.8}$$

From the boundary conditions, $U(x,s)$ is an even function in x and we may conveniently write the solution as

$$U(x,s) = \frac{1}{s} + A\cosh(qx). \tag{12.11.9}$$

From Equation 12.11.6,

$$qA\sinh(qL) + \frac{h}{s} + hA\cosh(qL) = 0, \tag{12.11.10}$$

and

$$U(x,s) = \frac{1}{s} - \frac{h\cosh(qx)}{s[q\sinh(qL) + h\cosh(qL)]}. \tag{12.11.11}$$

The inverse of $U(x,s)$ consists of two terms. The first term is simply unity. We will invert the second term by contour integration.

We begin by examining the nature and location of the singularities in the second term. Using the product formulas for the hyperbolic cosine and sine functions, the second term equals

$$\frac{h\left(1+\frac{4q^2x^2}{\pi^2}\right)\left(1+\frac{4q^2x^2}{9\pi^2}\right)\cdots}{s\left[q^2L\left(1+\frac{q^2L^2}{\pi^2}\right)\left(1+\frac{q^2L^2}{4\pi^2}\right)\cdots + h\left(1+\frac{4q^2L^2}{\pi^2}\right)\left(1+\frac{4q^2L^2}{9\pi^2}\right)\cdots\right]}. \tag{12.11.12}$$

Because $q^2 = s/a^2$, Equation 12.11.12 shows that we do not have any $\sqrt{s}$ in the transform and we need not concern ourselves with branch points and cuts. Furthermore, we have only simple poles: one located at $s = 0$ and the others where

$$q\sinh(qL) + h\cosh(qL) = 0. \tag{12.11.13}$$

If we set $q = i\lambda$, Equation 12.11.13 becomes

$$h\cos(\lambda L) - \lambda\sin(\lambda L) = 0, \qquad \text{or} \qquad \lambda L\tan(\lambda L) = hL. \tag{12.11.14}$$

From Bromwich's integral,

$$\mathcal{L}^{-1}\left\{\frac{h\cosh(qx)}{s[q\sinh(qL) + h\cosh(qL)]}\right\} = \frac{1}{2\pi i}\oint_C \frac{h\cosh(qx)e^{tz}}{z[q\sinh(qL) + h\cosh(qL)]}\,dz, \tag{12.11.15}$$

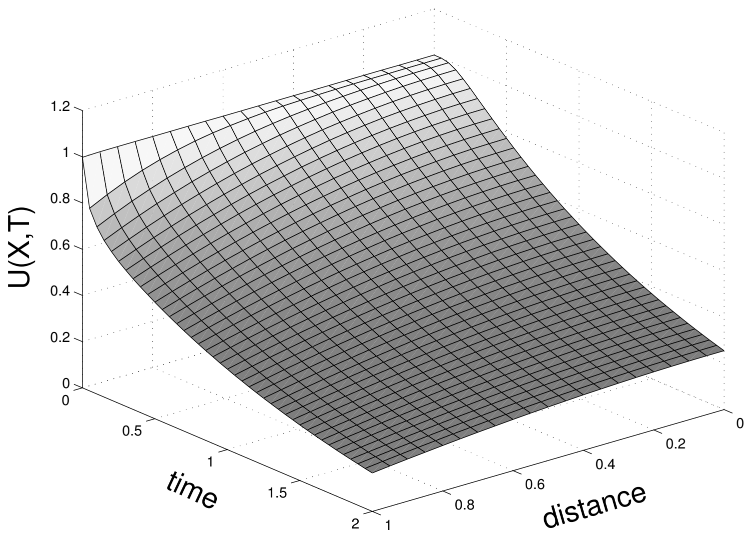


Figure 12.11.1: The temperature within the portion of a slab $0 < x/L < 1$ at various times a^2t/L^2 if the faces of the slab radiate to free space at temperature zero and the slab initially has the temperature 1. The parameter $hL = 1$.

where $q = z^{1/2}/a$ and the closed contour C consists of Bromwich's contour plus a semicircle of infinite radius in the left half of the z-plane. The residue at $z = 0$ is 1 while at $z_n = -a^2\lambda_n^2$,

$$\operatorname{Res}\left\{\frac{h\cosh(qx)e^{tz}}{z[q\sinh(qL)+h\cosh(qL)]}; z_n\right\} = \lim_{z\to z_n}\frac{h(z+a^2\lambda_n^2)\cosh(qx)e^{tz}}{z[q\sinh(qL)+h\cosh(qL)]} \quad \textbf{(12.11.16)}$$

$$= \lim_{z\to z_n}\frac{h\cosh(qx)e^{tz}}{z[(1+hL)\sinh(qL)+qL\cosh(qL)]/(2a^2q)} \quad \textbf{(12.11.17)}$$

$$= \frac{2ha^2\lambda_n i\cosh(i\lambda_n x)\exp(-\lambda_n^2a^2t)}{(-a^2\lambda_n^2)[(1+hL)i\sin(\lambda_nL)+i\lambda_nL\cos(\lambda_nL)]} \quad \textbf{(12.11.18)}$$

$$= -\frac{2h\cos(\lambda_nx)\exp(-a^2\lambda_n^2t)}{\lambda_n[(1+hL)\sin(\lambda_nL)+\lambda_nL\cos(\lambda_nL)]}. \quad \textbf{(12.11.19)}$$

Therefore, the inversion of $U(x,s)$ is

$$u(x,t) = 1 - \left\{1 - 2h\sum_{n=1}^{\infty}\frac{\cos(\lambda_nx)\exp(-a^2\lambda_n^2t)}{\lambda_n[(1+hL)\sin(\lambda_nL)+\lambda_nL\cos(\lambda_nL)]}\right\}, \quad \textbf{(12.11.20)}$$

or

$$u(x,t) = 2h\sum_{n=1}^{\infty}\frac{\cos(\lambda_nx)\exp(-a^2\lambda_n^2t)}{\lambda_n[(1+hL)\sin(\lambda_nL)+\lambda_nL\cos(\lambda_nL)]}. \quad \textbf{(12.11.21)}$$

We can further simplify Equation 12.11.21 by using $h/\lambda_n = \tan(\lambda_nL)$. This yields $hL = \lambda_nL\tan(\lambda_nL)$. Substituting these relationships into Equation 12.11.21 and simplifying,

$$u(x,t) = 2\sum_{n=1}^{\infty}\frac{\sin(\lambda_nL)\cos(\lambda_nx)\exp(-a^2\lambda_n^2t)}{\lambda_nL+\sin(\lambda_nL)\cos(\lambda_nL)}. \quad \textbf{(12.11.22)}$$

Figure 12.11.1 illustrates Equation 12.11.23. It was created using the MATLAB script

```
clear
hL = 1; m = 0; M = 100; dx = 0.05; dt = 0.05;
% create initial guess at zero_n
zero = zeros(length(M));
for n = 1:10000
  k1 = 0.1*n; k2 = 0.1*(n+1);
  prod = k1 * tan(k1); y1 = hL - prod; y2 = hL - k2 * tan(k2);
  if (y1*y2 <= 0 & prod < 2 & m < M) m = m+1; zero(m) = k1; end;
end;
% use Newton-Raphson method to improve values of zero_n
for n = 1:M; for k = 1:10
  f = hL - zero(n) * tan(zero(n));
  fp = - tan(zero(n)) - zero(n) * sec(zero(n))^2;
  zero(n) = zero(n) - f / fp;
end; end;
% compute Fourier coefficients
for m = 1:M
  a(m) = 2 * sin(zero(m)) / (zero(m) + sin(zero(m))*cos(zero(m)));
end
% compute grid and initialize solution
X = [0:dx:1]; T = [0:dt:2];
u = zeros(length(T),length(X));
XX = repmat(X,[length(T) 1]); TT = repmat(T',[1 length(X)]);
% compute solution from Equation 12.11.22
for m = 1:M
  u = u + a(m) * cos(zero(m)*XX) .* exp(-zero(m)*zero(m)*TT);
end
surf(XX,TT,u)
xlabel('distance','Fontsize',20); ylabel('time','Fontsize',20)
zlabel('U(X,T)','Fontsize',20)
```
□

• Example 12.11.2: Heat dissipation in disc brakes

Disc brakes consist of two blocks of frictional material known as pads that press against each side of a rotating annulus, usually made of a ferrous material. In this problem we determine the transient temperatures reached in a disc brake during a single brake application.[48] If we ignore the errors introduced by replacing the cylindrical portion of the drum by a rectangular plate, we can model our disc brakes as a one-dimensional solid, which friction heats at both ends. Assuming symmetry about $x = 0$, the boundary condition there is $u_x(0, t) = 0$. To model the heat flux from the pads, we assume a uniform disc deceleration that generates heat from the frictional surfaces at the rate $N(1 - Mt)$, where M and N are experimentally determined constants.

If $u(x, t)$, κ, and a^2 denote the temperature, thermal conductivity, and diffusivity of the rotating annulus, respectively, then the heat equation is

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{12.11.23}$$

[48] From Newcomb, T. P., 1958: The flow of heat in a parallel-faced infinite solid. *Brit. J. Appl. Phys.*, **9**, 370–372. See also Newcomb, T. P., 1958/59: Transient temperatures in brake drums and linings. *Proc. Inst. Mech. Eng., Auto. Div.*, 227–237; Newcomb, T. P., 1959: Transient temperatures attained in disk brakes. *Brit. J. Appl. Phys.*, **10**, 339–340.

with the boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = 0, \quad \kappa \frac{\partial u(L,t)}{\partial x} = N(1 - Mt), \qquad 0 < t. \tag{12.11.24}$$

The boundary condition at $x = L$ gives the frictional heating of the disc pads.

Introducing the Laplace transform of $u(x,t)$, defined as

$$U(x,s) = \int_0^\infty u(x,t)e^{-st}\,dt, \tag{12.11.25}$$

the equation to be solved becomes

$$\frac{d^2U}{dx^2} - \frac{s}{a^2}U = 0, \tag{12.11.26}$$

subject to the boundary conditions that

$$\frac{dU(0,s)}{dx} = 0, \quad \text{and} \quad \frac{dU(L,s)}{dx} = \frac{N}{\kappa}\left(\frac{1}{s} - \frac{M}{s^2}\right). \tag{12.11.27}$$

The solution of Equation 12.11.26 is

$$U(x,s) = A\cosh(qx) + B\sinh(qx), \tag{12.11.28}$$

where $q = s^{1/2}/a$. Using the boundary conditions, the solution becomes

$$U(x,s) = \frac{N}{\kappa}\left(\frac{1}{s} - \frac{M}{s^2}\right)\frac{\cosh(qx)}{q\sinh(qL)}. \tag{12.11.29}$$

It now remains to invert the transform, Equation 12.11.29. We will invert $\cosh(qx)/[sq\sinh(qL)]$; the inversion of the second term follows by analog.

Our first concern is the presence of $s^{1/2}$ because this is a multivalued function. However, when we replace the hyperbolic cosine and sine functions with their Taylor expansions, $\cosh(qx)/[sq\sinh(qL)]$ contains only powers of s and is, in fact, a single-valued function.

From Bromwich's integral,

$$\mathcal{L}^{-1}\left[\frac{\cosh(qx)}{sq\sinh(qL)}\right] = \frac{1}{2\pi i}\int_{c-\infty i}^{c+\infty i} \frac{\cosh(qx)e^{tz}}{zq\sinh(qL)}\,dz, \tag{12.11.30}$$

where $q = z^{1/2}/a$. Just as in the previous example, we replace the hyperbolic cosine and sine with their product expansion to determine the nature of the singularities. The point $z = 0$ is a second-order pole. The remaining poles are located where $z_n^{1/2}L/a = n\pi i$, or $z_n = -n^2\pi^2a^2/L^2$, where $n = 1, 2, 3, \ldots$. We have chosen the positive sign because $z^{1/2}$ must be single-valued; if we had chosen the negative sign the answer would have been the same. Our expansion also shows that the poles are simple.

Having classified the poles, we now close Bromwich's contour, which lies slightly to the right of the imaginary axis, with an infinite semicircle in the left half-plane, and use the residue theorem. The values of the residues are

$$\operatorname{Res}\left[\frac{\cosh(qx)e^{tz}}{zq\sinh(qL)};0\right] = \frac{1}{1!}\lim_{z\to0}\frac{d}{dz}\left\{\frac{(z-0)^2\cosh(qx)e^{tz}}{zq\sinh(qL)}\right\} \tag{12.11.31}$$

$$= \lim_{z\to0}\frac{d}{dz}\left\{\frac{z\cosh(qx)e^{tz}}{q\sinh(qL)}\right\} \tag{12.11.32}$$

$$= \frac{a^2}{L}\lim_{z\to0}\frac{d}{dz}\left\{\frac{z\left[1+\frac{zx^2}{2!a^2}+\cdots\right]\left[1+tz+\frac{t^2z^2}{2!}+\cdots\right]}{z+\frac{L^2z^2}{3!a^2}+\cdots}\right\} \tag{12.11.33}$$

$$= \frac{a^2}{L}\lim_{z\to0}\frac{d}{dz}\left\{1+tz+\frac{zx^2}{2a^2}-\frac{zL^2}{3!a^2}+\cdots\right\} \tag{12.11.34}$$

$$= \frac{a^2}{L}\left\{t+\frac{x^2}{2a^2}-\frac{L^2}{6a^2}\right\}, \tag{12.11.35}$$

and

$$\operatorname{Res}\left[\frac{\cosh(qx)e^{tz}}{zq\sinh(qL)};z_n\right] = \left[\lim_{z\to z_n}\frac{\cosh(qx)}{zq}e^{tz}\right]\left[\lim_{z\to z_n}\frac{z-z_n}{\sinh(qL)}\right] \tag{12.11.36}$$

$$= \lim_{z\to z_n}\frac{\cosh(qx)e^{tz}}{zq\cosh(qL)L/(2a^2q)} \tag{12.11.37}$$

$$= \frac{\cosh(n\pi xi/L)\exp(-n^2\pi^2a^2t/L^2)}{(-n^2\pi^2a^2/L^2)\cosh(n\pi i)L/(2a^2)} \tag{12.11.38}$$

$$= -\frac{2L(-1)^n}{n^2\pi^2}\cos(n\pi x/L)e^{-n^2\pi^2a^2t/L^2}. \tag{12.11.39}$$

When we sum all of the residues from both inversions, the solution is

$$\begin{aligned} u(x,t) = {} & \frac{a^2N}{\kappa L}\left\{t+\frac{x^2}{2a^2}-\frac{L^2}{6a^2}\right\} - \frac{2LN}{\kappa\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(n\pi x/L)e^{-n^2\pi^2a^2t/L^2} \\ & - \frac{a^2NM}{\kappa L}\left\{\frac{t^2}{2}+\frac{tx^2}{2a^2}-\frac{tL^2}{6a^2}+\frac{x^4}{24a^4}-\frac{x^2L^2}{12a^4}+\frac{7L^4}{360a^4}\right\} \\ & - \frac{2L^3NM}{a^2\kappa\pi^4}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^4}\cos(n\pi x/L)e^{-n^2\pi^2a^2t/L^2}. \end{aligned} \tag{12.11.40}$$

Figure 12.11.2 shows the temperature in the brake lining at various places within the lining $[x' = x/L]$ if $a^2 = 3.3\times10^{-3}$ cm^2/sec, $\kappa = 1.8\times10^{-3}$ cal/(cm sec°C), $L = 0.48$ cm, and $N = 1.96$ cal/(cm^2 sec). Initially the frictional heating results in an increase in the disc brake's temperature. As time increases, the heating rate decreases and radiative cooling becomes sufficiently large that the temperature begins to fall. □

• Example 12.11.3

In the previous example we showed that Laplace transforms are particularly useful when the boundary conditions are time dependent. Consider now the case when one of the boundaries is moving.

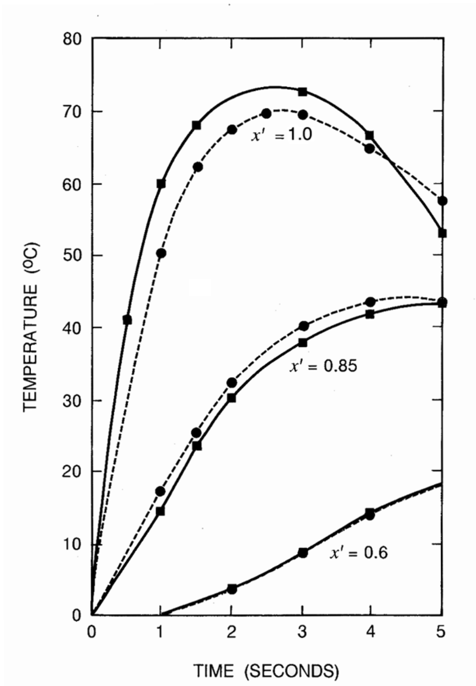


Figure 12.11.2: Typical curves of transient temperature at different locations in a brake lining. Circles denote computed values while squares are experimental measurements. (From Newcomb, T. P., 1958: The flow of heat in a parallel-faced infinite solid. *Brit. J. Appl. Phys.*, **9**, 372 with permission.)

We wish to solve the heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad \beta t < x < \infty, \quad 0 < t, \tag{12.11.41}$$

subject to the boundary conditions

$$u(x,t)\big|_{x=\beta t} = f(t), \quad \text{and} \quad \lim_{x\to\infty} u(x,t) \to 0, \quad 0 < t, \tag{12.11.42}$$

and the initial condition

$$u(x,0) = 0, \qquad 0 < x < \infty. \tag{12.11.43}$$

This type of problem arises in combustion problems where the boundary moves due to the burning of the fuel.

We begin by introducing the coordinate $\eta = x - \beta t$. Then the problem can be reformulated as

$$\frac{\partial u}{\partial t} - \beta \frac{\partial u}{\partial \eta} = a^2 \frac{\partial^2 u}{\partial \eta^2}, \quad 0 < \eta < \infty, \quad 0 < t, \tag{12.11.44}$$

subject to the boundary conditions

$$u(0,t) = f(t), \quad \lim_{\eta\to\infty} u(\eta,t) \to 0, \qquad 0 < t, \tag{12.11.45}$$

and the initial condition

$$u(\eta,0) = 0, \qquad 0 < \eta < \infty. \tag{12.11.46}$$

Taking the Laplace transform of Equation 12.11.44, we have that

$$\frac{d^2U(\eta,s)}{d\eta^2}+\frac{\beta}{a^2}\frac{dU(\eta,s)}{d\eta}-\frac{s}{a^2}U(\eta,s)=0, \tag{12.11.47}$$

with

$$U(0,s)=F(s), \quad \text{and} \quad \lim_{\eta\to\infty}U(\eta,s)\to 0. \tag{12.11.48}$$

The solution to Equation 12.11.47 and Equation 12.11.48 is

$$U(\eta,s)=F(s)\exp\left(-\frac{\beta\eta}{2a^2}-\frac{\eta}{a}\sqrt{s+\frac{\beta^2}{4a^2}}\right). \tag{12.11.49}$$

Because

$$\mathcal{L}\left[\Phi(\eta,t)\right]=\exp\left(-\frac{\eta}{a}\sqrt{s+\frac{\beta^2}{4a^2}}\right), \tag{12.11.50}$$

where

$$\Phi(\eta,t)=\tfrac{1}{2}\left[e^{-\beta\eta/2a^2}\operatorname{erfc}\left(\frac{\eta}{2a\sqrt{t}}-\frac{\beta\sqrt{t}}{2a}\right)+e^{\beta\eta/2a^2}\operatorname{erfc}\left(\frac{\eta}{2a\sqrt{t}}+\frac{\beta\sqrt{t}}{2a}\right)\right], \tag{12.11.51}$$

and

$$\operatorname{erfc}(x)=1-\frac{2}{\sqrt{\pi}}\int_0^x e^{-\eta^2}\,d\eta, \tag{12.11.52}$$

we have by the convolution theorem that

$$u(\eta,t)=e^{-\beta\eta/2a^2}\int_0^t f(t-\tau)\Phi(\eta,\tau)\,d\tau, \tag{12.11.53}$$

or

$$u(x,t)=e^{-\beta(x-\beta t)/2a^2}\int_0^t f(t-\tau)\Phi(x-\beta\tau,\tau)\,d\tau. \tag{12.11.54}$$

Problems

1. Solve

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}-a^2(u-T_0), \qquad 0<x<1, \quad 0<t,$$

subject to the boundary conditions $u_x(0,t)=u_x(1,t)=0$, $0<t$, and the initial condition $u(x,0)=0$, $0<x<1$.

2. Solve

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}, \qquad 0<x<1, \quad 0<t,$$

subject to the boundary conditions $u_x(0,t)=0$, $u(1,t)=t$, $0<t$, and the initial condition $u(x,0)=0$, $0<x<1$.

3. Solve

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}, \qquad 0<x<1, \quad 0<t,$$

subject to the boundary conditions $u(0,t) = 0$, $u(1,t) = 1$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < 1$.

4. Solve

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad -\tfrac{1}{2} < x < \tfrac{1}{2}, \quad 0 \leq t,$$

subject to the boundary conditions $u_x\left(-\frac{1}{2}, t\right) = 0$, $u_x\left(\frac{1}{2}, t\right) = \delta(t)$, $0 \leq t$, and the initial condition $u(x,0) = 0$, $-\frac{1}{2} < x < \frac{1}{2}$.

5. Solve

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 1, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions $u(0,t) = u(1,t) = 0$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < 1$.

6. Solve[49]

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions

$$u(0,t) = 1, \quad \lim_{x\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = 0$, $0 < x < \infty$. Hint: Use tables to invert the Laplace transform.

7. Solve

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = 1, \quad \lim_{x\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = 0$, $0 < x < \infty$. Hint: Use tables to invert the Laplace transform.

8. Solve

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions

$$u(0,t) = 1, \quad \lim_{x\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = e^{-x}$, $0 < x < \infty$. Hint: Use tables to invert the Laplace transform.

---

[49] If $u(x,t)$ denotes the Eulerian velocity of a viscous fluid in the half space $x > 0$ and parallel to the wall located at $x = 0$, then this problem was first solved by Stokes, G. G., 1850: On the effect of the internal friction of fluids on the motions of pendulums. *Proc. Cambridge Philos. Soc.*, **9, Part II**, [8]–[106].

9. Solve

$$\frac{\partial u}{\partial t} = a^2 \left[\frac{\partial^2 u}{\partial x^2} + (1+\delta)\frac{\partial u}{\partial x} + \delta u\right], \qquad 0 < x < \infty, \quad 0 < t,$$

where δ is a constant, subject to the boundary conditions

$$u(0,t) = u_0, \quad \lim_{x\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = 0$, $0 < x < \infty$. Note that

$$\mathcal{L}^{-1}\left[\frac{1}{s}\exp\left(-2\alpha\sqrt{s+\beta^2}\,\right)\right] = \tfrac{1}{2}e^{2\alpha\beta}\text{erfc}\left(\frac{\alpha}{\sqrt{t}} + \beta\sqrt{t}\right) + \tfrac{1}{2}e^{-2\alpha\beta}\text{erfc}\left(\frac{\alpha}{\sqrt{t}} - \beta\sqrt{t}\right),$$

where $\text{erfc}(\cdot)$ is the complementary error function.

10. During their modeling of a chemical reaction with a back reaction, Agmon et al.[50] solved

$$\frac{\partial u}{\partial t} = a^2\frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions

$$\kappa_d + a^2 u_x(0,t) + a^2\kappa_d \int_0^t u_x(0,\tau)\,d\tau = \kappa_r u(0,t),$$

$$\lim_{x\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = 0$, $0 < x < \infty$, where κ_d and κ_r denote the intrinsic dissociation and recombination rate coefficients, respectively. What should they have found?

11. Solve[51]

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - \beta u, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions

$$\rho u(0,t) - u_x(0,t) = e^{(\sigma^2-\beta)t}, \quad \lim_{x\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = 0$, $0 < x < \infty$, where β, ρ, and σ are constants and $\sigma \neq \rho$.

12. Solve

$$\frac{\partial u}{\partial t} = a^2\frac{\partial^2 u}{\partial x^2} + Ae^{-kx}, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions

$$\frac{\partial u(0,t)}{\partial x} = 0, \quad \lim_{x\to\infty} u(x,t) = u_0, \qquad 0 < t,$$

---

[50] Agmon, N., E. Pines, and D. Huppert, 1988: Geminate recombination in proton-transfer reactions. II. Comparison of diffusional and kinetic schemes. *J. Chem. Phys.*, **88**, 5631–5638.

[51] Saidel, G. M., E. D. Morris, and G. M. Chisolm, 1987: Transport of macromolecules in arterial wall *in vivo*: A mathematical model and analytic solutions. *Bull. Math. Biol.*, **49**, 153–169.

and the initial condition $u(x,0) = u_0$, $0 < x < \infty$.

13. Solve

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} - P, \qquad 0 < x < L, \quad 0 < t,$$

subject to the boundary conditions $u(0,t) = t$, $u(L,t) = 0$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < L$.

14. Solve

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + ku, \qquad 0 < x < L, \quad 0 < k, t,$$

subject to the boundary conditions $u(0,t) = u(L,t) = T_0$, $0 < t$, and the initial condition $u(x,0) = T_0$, $0 < x < L$.

15. An electric fuse protects electrical devices by using resistance heating to melt an enclosed wire when excessive current passes through it. A knowledge of the distribution of temperature along the wire is important in the design of the fuse. If the temperature rises to the melting point only over a small interval of the element, the melt will produce a small gap, resulting in an unnecessary prolongation of the fault and a considerable release of energy. Therefore, the desirable temperature distribution should melt most of the wire. For this reason, Guile and Carne[52] solved the heat conduction equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + q(1 + \alpha u), \qquad -L < x < L, \quad 0 < t,$$

to understand the temperature structure within the fuse just before meltdown. The second term on the right side of the heat conduction equation gives the resistance heating, which is assumed to vary linearly with temperature. If the terminals at $x = \pm L$ remain at a constant temperature, which we can take to be zero, the boundary conditions are $u(-L,t) = u(L,t) = 0$, $0 < t$. The initial condition is $u(x,0) = 0$, $-L < x < L$. Find the temperature field as a function of the parameters a, q, and α.

16. Solve[53]

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r}, \qquad 0 \le r < 1, \quad 0 < t,$$

subject to the boundary conditions

$$\lim_{r\to 0} |u(r,t)| < \infty, \qquad \frac{\partial u(1,t)}{\partial r} = 1, \qquad 0 < t,$$

and the initial condition $u(r,0) = 0$, $0 \le r < 1$. Hint: Use the new dependent variable $v(r,t) = ru(r,t)$.

17. Solve[54]

$$\frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} \right) + q(t) = \frac{a^2}{r}\frac{\partial^2 (ru)}{\partial r^2} + q(t), \qquad b < r < \infty, \quad 0 < t,$$

---

[52] Guile, A. E., and E. B. Carne, 1954: An analysis of an analogue solution applied to the heat conduction problem in a cartridge fuse. *AIEE Trans., Part 1*, **72**, 861–868.

[53] See Reismann, H., 1962: Temperature distribution in a spinning sphere during atmospheric entry. *J. Aerosp. Sci.*, **29**, 151–159.

[54] See Frisch, H. L, and F. C. Collins, 1952: Diffusional processes in the growth of aerosol particles. *J. Chem. Phys.*, **20**, 1797–1803.

subject to the boundary conditions

$$\frac{\partial u(b,t)}{\partial r} = u(b,t), \qquad \lim_{r\to\infty} u(r,t) = u_0 + \int_0^t q(\tau)\,d\tau, \qquad 0 < t,$$

and the initial condition $u(r,0) = u_0$, $b < r < \infty$.

18. Consider[55] a viscous fluid located between two fixed walls $x = \pm L$. At $x = 0$ we introduce a thin, infinitely long rigid barrier of mass m per unit area and let it fall under the force of gravity, which points in the direction of positive x. We wish to find the velocity of the fluid $u(x,t)$. The fluid is governed by the partial differential equation

$$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

subject to the boundary conditions

$$u(L,t) = 0, \qquad \frac{\partial u(0,t)}{\partial t} - \frac{2\mu}{m}\frac{\partial u(0,t)}{\partial x} = g, \qquad 0 < t,$$

and the initial condition $u(x,0) = 0$, $0 < x < L$.

19. Consider[56] a viscous fluid located between two fixed walls $x = \pm L$. At $x = 0$ we introduce a thin, infinitely long rigid barrier of mass m per unit area. The barrier is acted upon by an elastic force in such a manner that it would vibrate with a frequency ω if the liquid were absent. We wish to find the barrier's deviation from equilibrium, $y(t)$. The fluid is governed by the partial differential equation

$$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t.$$

The boundary conditions are

$$u(L,t) = m\frac{d^2 y}{dt^2} - 2\mu \frac{\partial u(0,t)}{\partial x} + m\omega^2 y = 0, \qquad \frac{dy}{dt} = u(0,t), \quad 0 < t,$$

and the initial conditions are $u(x,0) = 0$, $0 < x < L$, and $y(0) = A$, $y'(0) = 0$.

20. Solve[57]

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

subject to the boundary conditions $u_x(0,t) = 0$, $a^2 u_x(L,t) + \alpha u(L,t) = F$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < L$.

---

[55] See Havelock, T. H., 1921: The solution of an integral equation occurring in certain problems of viscous fluid motion. *Philos. Mag., Ser. 6*, **42**, 620–628.

[56] See Havelock, T. H., 1921: On the decay of oscillation of a solid body in a viscous fluid. *Philos. Mag., Ser. 6*, **42**, 628–634.

[57] See McCarthy, T. A., and H. J. Goldsmid, 1970: Electro-deposited copper in bismuth telluride. *J. Phys. D*, **3**, 697–706.

21. Solve

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 \le x < 1, \quad 0 \le t,$$

subject to the boundary conditions

$$u(0,t) = 0, \quad 3a\left[\frac{\partial u(1,t)}{\partial x} - u(1,t)\right] + \frac{\partial u(1,t)}{\partial t} = \delta(t), \quad 0 \le t,$$

and the initial condition $u(x,0) = 0$, $0 \le x < 1$.

22. Solve[58] the partial differential equation

$$\frac{\partial u}{\partial t} + V\frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

where V is a constant, subject to the boundary conditions

$$u(0,t) = 1, \quad u_x(1,t) = 0, \qquad 0 < t,$$

and the initial condition $u(x,0) = 0$, $0 < x < 1$.

23. Solve[59] the partial differential equation

$$\frac{\partial^2 u}{\partial x \partial t} + a\frac{\partial u}{\partial t} + b\frac{\partial u}{\partial x} = 0, \qquad 0 < x < \infty, \quad 0 < a, b, t,$$

subject to the boundary conditions

$$u(0,t) = 1, \quad \lim_{x\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $u_x(x,0) + au(x,0) = 0$, $0 < x < \infty$. To invert the Laplace transform, you may want to review Problem 52 at the end of Section 12.3.

24. Solve

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) - \frac{\partial u}{\partial t} = \delta(t), \qquad 0 \le r < a, \quad 0 \le t,$$

subject to the boundary conditions

$$\lim_{r\to 0} |u(r,t)| < \infty, \quad u(a,t) = 0, \qquad 0 \le t,$$

and the initial condition $u(r,0) = 0$, $0 \le r < a$, where $\delta(t)$ is the Dirac delta function. Note that $J_n(iz) = i^n I_n(z)$ and $I_n(iz) = i^n J_n(z)$ for all complex z.

25. Solve

$$\frac{\partial u}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + H(t), \qquad 0 \le r < a, \quad 0 < t,$$

---

[58] See Yoo, H., and E.-T. Pak, 1996: Analytical solutions to a one-dimensional finite-domain model for stratified thermal storage tanks. *Sol. Energy*, **56**, 315–322.

[59] See Liaw, C. H., J. S. P. Wang, R. A. Greenhorn, and K. C. Chao, 1979: Kinetics of fixed-bed absorption: A new solution. *AICHE J.*, **25**, 376–381.

subject to the boundary conditions

$$\lim_{r\to 0} |u(r,t)| < \infty, \quad u(a,t) = 0, \qquad 0 < t,$$

and the initial condition $u(r,0) = 0$, $0 \le r < a$. Note that $J_n(iz) = i^n I_n(z)$ and $I_n(iz) = i^n J_n(z)$ for all complex z.

26. Solve

$$\frac{\partial u}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right), \qquad 0 \le r < a, \quad 0 < t,$$

subject to the boundary conditions

$$\lim_{r\to 0} |u(r,t)| < \infty, \quad u(a,t) = e^{-t/\tau_0}, \qquad 0 < t,$$

and the initial condition $u(r,0) = 1$, $0 \le r < a$. Note that $J_n(iz) = i^n I_n(z)$ and $I_n(iz) = i^n J_n(z)$ for all complex z.

27. Solve

$$\frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right), \qquad 0 \le r < b, \quad 0 < t,$$

subject to the boundary conditions

$$\lim_{r\to 0} |u(r,t)| < \infty, \quad u(b,t) = kt, \qquad 0 < t,$$

and the initial condition $u(r,0) = 0$, $0 \le r < b$. Note that $J_n(iz) = i^n I_n(z)$ and $I_n(iz) = i^n J_n(z)$ for all complex z.

28. Solve the nonhomogeneous heat equation for the spherical shell[60]

$$\frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{A}{r^4}\right), \qquad \alpha < r < \beta, \quad 0 < t,$$

subject to the boundary conditions

$$\frac{\partial u(\alpha,t)}{\partial r} = u(\beta,t) = 0, \qquad 0 < t,$$

and the initial condition $u(r,0) = 0$, $\alpha < r < \beta$.

Step 1: By introducing $v(r,t) = r\,u(r,t)$, show that the problem simplifies to

$$\frac{\partial v}{\partial t} = a^2\left(\frac{\partial^2 v}{\partial r^2} + \frac{A}{r^3}\right), \qquad \alpha < r < \beta, \quad 0 < t,$$

subject to the boundary conditions

$$\frac{\partial v(\alpha,t)}{\partial r} - \frac{v(\alpha,t)}{\alpha} = v(\beta,t) = 0, \qquad 0 < t,$$

[60] See Malkovich, R. Sh., 1977: Heating of a spherical shell by a radial current. *Sov. Phys. Tech. Phys.*, **22**, 636.

and the initial condition

$$v(r,0) = 0, \qquad \alpha < r < \beta.$$

Step 2: Using Laplace transforms and variation of parameters, show that the Laplace transform of $u(r,t)$ is

$$U(r,s) = \frac{A}{srq}\left\{\frac{\sinh[q(\beta - r)]}{\alpha q\,\cosh(q\ell\,) + \sinh(q\ell\,)}\int_0^\ell \frac{\alpha q\,\cosh(q\eta) + \sinh(q\eta)}{(\alpha+\eta)^3}\,d\eta - \int_0^{\beta-r}\frac{\sinh(q\eta)}{(r+\eta)^3}\,d\eta\right\},$$

where $q = \sqrt{s}/a$, and $\ell = \beta - \alpha$.

Step 3: Take the inverse of $U(r,s)$ and show that

$$u(r,t) = A\left\{\left(\frac{1}{r} - \frac{1}{\beta}\right)\left[\frac{1}{\alpha} - \frac{1}{2}\left(\frac{1}{r} + \frac{1}{\beta}\right)\right] - \frac{2\alpha^2}{r\ell^2}\sum_{n=0}^{\infty}\frac{\sin[\gamma_n(\beta - r)]\exp(-a^2\gamma_n^2 t)}{\sin^2(\gamma_n\ell\,)(\beta + \alpha^2\ell\gamma_n^2)}\int_0^1\frac{\sin(\gamma_n\ell\eta)}{(\delta-\eta)^3}\,d\eta\right\},$$

where γ_n is the nth root of $\alpha\gamma + \tan(\ell\gamma) = 0$, and $\delta = 1 + \alpha/\ell$.

12.12 THE SUPERPOSITION INTEGRAL AND THE HEAT EQUATION

In our study of Laplace transforms, we showed that we can construct solutions to ordinary differential equations with a general forcing $f(t)$ by first finding the solution to a similar problem where the forcing equals Heaviside's step function. Then we can write the general solution in terms of a superposition integral according to Duhamel's theorem. In this section we show that similar considerations hold in solving the heat equation with time-dependent boundary conditions or forcings.

Let us solve the heat condition problem

$$\frac{\partial u}{\partial t} = a^2\frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{12.12.1}$$

with the boundary conditions

$$u(0,t) = 0, \quad u(L,t) = f(t), \quad 0 < t, \tag{12.12.2}$$

and the initial condition

$$u(x,0) = 0, \quad 0 < x < L. \tag{12.12.3}$$

The solution of Equation 12.12.1 through Equation 12.12.3 is difficult because of the time-dependent boundary condition. Instead of solving this system directly, let us solve the easier problem

$$\frac{\partial A}{\partial t} = a^2\frac{\partial^2 A}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{12.12.4}$$

with the boundary conditions

$$A(0,t) = 0, \quad A(L,t) = 1, \quad 0 < t, \tag{12.12.5}$$

and the initial condition

$$A(x,0) = 0, \quad 0 < x < L. \tag{12.12.6}$$

Separation of variables yields the solution

$$A(x,t) = \frac{x}{L} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin\left(\frac{n\pi x}{L}\right)\exp\left(-\frac{a^2n^2\pi^2 t}{L^2}\right). \tag{12.12.7}$$

Consider the following case. Suppose that we maintain the temperature at zero at the end $x = L$ until $t = \tau_1$ and then raise it to the value of unity. The resulting temperature distribution equals zero everywhere when $t < \tau_1$ and equals $A(x, t-\tau_1)$ for $t > \tau_1$. We have merely shifted our time axis so that the initial condition occurs at $t = \tau_1$.

Consider an analogous, but more complicated, situation of the temperature at the end position $x = L$ held at $f(0)$ from $t = 0$ to $t = \tau_1$ at which time we abruptly change it by the amount $f(\tau_1) - f(0)$ to the value $f(\tau_1)$. This temperature remains until $t = \tau_2$ when we again abruptly change it by an amount $f(\tau_2) - f(\tau_1)$. We can imagine this process continuing up to the instant $t = \tau_n$. Because of linear superposition, the temperature distribution at any given time equals the sum of these temperature increments:

$$\begin{aligned} u(x,t) = f(0)A(x,t) &+ [f(\tau_1) - f(0)]A(x, t-\tau_1) + [f(\tau_2) - f(\tau_1)]A(x, t-\tau_2) \\ &+ \cdots + [f(\tau_n) - f(\tau_{n-1})]A(x, t-\tau_n), \end{aligned} \tag{12.12.8}$$

where τ_n is the time of the most recent temperature change. If we write

$$\Delta f_k = f(\tau_k) - f(\tau_{k-1}), \quad \text{and} \quad \Delta\tau_k = \tau_k - \tau_{k-1}, \tag{12.12.9}$$

Equation 12.12.8 becomes

$$u(x,t) = f(0)A(x,t) + \sum_{k=1}^{n} A(x, t-\tau_k)\frac{\Delta f_k}{\Delta\tau_k}\Delta\tau_k. \tag{12.12.10}$$

Consequently, in the limit of $\Delta\tau_k \to 0$, Equation 12.12.10 becomes

$$u(x,t) = f(0)A(x,t) + \int_0^t A(x, t-\tau)f'(\tau)\,d\tau, \tag{12.12.11}$$

assuming that $f(t)$ is differentiable. Equation 12.12.11 is the *superposition integral.* We can obtain an alternative form by integration by parts:

$$u(x,t) = f(t)A(x,0) - \int_0^t f(\tau)\frac{\partial A(x, t-\tau)}{\partial\tau}\,d\tau, \tag{12.12.12}$$

or

$$u(x,t) = f(t)A(x,0) + \int_0^t f(\tau)\frac{\partial A(x, t-\tau)}{\partial t}\,d\tau, \tag{12.12.13}$$

because

$$\frac{\partial A(x, t-\tau)}{\partial\tau} = -\frac{\partial A(x, t-\tau)}{\partial t}. \tag{12.12.14}$$

To illustrate[61] the superposition integral, suppose $f(t) = t$. Then, by Equation 12.12.11,

$$u(x,t) = \int_0^t \left\{ \frac{x}{L} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{n\pi x}{L}\right) \exp\left[-\frac{a^2 n^2 \pi^2}{L^2}(t-\tau) \right] \right\} d\tau \tag{12.12.15}$$

$$= \frac{xt}{L} + \frac{2L^2}{a^2\pi^3} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^3} \sin\left(\frac{n\pi x}{L}\right) \left[1 - \exp\left(-\frac{a^2 n^2 \pi^2 t}{L^2} \right) \right]. \tag{12.12.16}$$

• Example 12.12.1: Temperature oscillations in a wall heated by an alternating current

In addition to finding solutions to heat conduction problems with time-dependent boundary conditions, we can also apply the superposition integral to the nonhomogeneous heat equation when the source depends on time. Jeglic[62] used this technique in obtaining the temperature distribution within a slab heated by alternating electric current. If we assume that the flat plate has a surface area A and depth L, then the heat equation for the plate when electrically heated by an alternating current of frequency ω is

$$\frac{\partial u}{\partial t} - a^2 \frac{\partial^2 u}{\partial x^2} = \frac{2q}{\rho C_p A L} \sin^2(\omega t), \qquad 0 < x < L, \quad 0 < t, \tag{12.12.17}$$

where q is the average heat rate caused by the current, ρ is the density, C_p is the specific heat at constant pressure, and a^2 is the diffusivity of the slab. We will assume that we insulated the inner wall so that

$$\frac{\partial u(0,t)}{\partial x} = 0, \qquad 0 < t, \tag{12.12.18}$$

while we allow the outer wall to radiatively cool to free space at the temperature of zero or

$$\kappa \frac{\partial u(L,t)}{\partial x} + h u(L,t) = 0, \qquad 0 < t, \tag{12.12.19}$$

where κ is the thermal conductivity and h is the heat transfer coefficient. The slab is initially at the temperature of zero or

$$u(x,0) = 0, \qquad 0 < x < L. \tag{12.12.20}$$

To solve the heat equation, we first solve the simpler problem of

$$\frac{\partial A}{\partial t} - a^2 \frac{\partial^2 A}{\partial x^2} = 1, \qquad 0 < x < L, \quad 0 < t, \tag{12.12.21}$$

---

[61] This occurs, for example, in McAfee, K. B., 1958: Stress-enhanced diffusion in glass. I. Glass under tension and compression. *J. Chem. Phys.*, **28**, 218–226. McAfee used an alternative method of guessing the solution.

[62] Jeglic, F. A., 1962: An analytical determination of temperature oscillations in a wall heated by alternating current. *NASA Tech. Note No. D–1286*. In a similar vein, Al-Nimr and Abdallah (Al-Nimr, M. A., and M. R. Abdallah, 1999: Thermal behavior of insulated electric wires producing pulsating signals. *Heat Transfer Eng.*, **20(4)**, 62–74) found the heat transfer with an insulated wire that carries an alternating current.

Table 12.12.1: The First Six Roots of the Equation $k_n \tan(k_n) = h^*$

h^*	k_1	k_2	k_3	k_4	k_5	k_6
0.001	0.03162	3.14191	6.28334	9.42488	12.56645	15.70803
0.002	0.04471	3.14223	6.28350	9.42499	12.56653	15.70809
0.005	0.07065	3.14318	6.28398	9.42531	12.56677	15.70828
0.010	0.09830	3.14477	6.28478	9.42584	12.56717	15.70860
0.020	0.14095	3.14795	6.28637	9.42690	12.56796	15.70924
0.050	0.22176	3.15743	6.29113	9.43008	12.57035	15.71115
0.100	0.31105	3.17310	6.29906	9.43538	12.57432	15.71433
0.200	0.43284	3.20393	6.31485	9.44595	12.58226	15.72068
0.500	0.65327	3.29231	6.36162	9.47748	12.60601	15.73972
1.000	0.86033	3.42562	6.43730	9.52933	12.64529	15.77128
2.000	1.07687	3.64360	6.57833	9.62956	12.72230	15.83361
5.000	1.31384	4.03357	6.90960	9.89275	12.93522	16.01066
10.000	1.42887	4.30580	7.22811	10.20026	13.21418	16.25336
20.000	1.49613	4.49148	7.49541	10.51167	13.54198	16.58640
∞	1.57080	4.71239	7.85399	10.99557	14.13717	17.27876

with the boundary conditions

$$\frac{\partial A(0,t)}{\partial x} = 0, \quad \kappa \frac{\partial A(L,t)}{\partial x} + hA(L,t) = 0, \qquad 0 < t, \tag{12.12.22}$$

and the initial condition

$$A(x,0) = 0, \qquad 0 < x < L. \tag{12.12.23}$$

The solution $A(x,t)$ is the *indicial admittance* because it is the response of a system to forcing by the step function $H(t)$.

We solve Equation 12.12.21 through Equation 12.12.23 by separation of variables. We begin by assuming that $A(x,t)$ consists of a steady-state solution $w(x)$ plus a transient solution $v(x,t)$, where

$$a^2 w''(x) = -1, \quad w'(0) = 0, \quad \kappa w'(L) + hw(L) = 0, \tag{12.12.24}$$

$$\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 v}{\partial x^2}, \quad \frac{\partial v(0,t)}{\partial x} = 0, \quad \kappa \frac{\partial v(L,t)}{\partial x} + hv(L,t) = 0, \tag{12.12.25}$$

and

$$v(x,0) = -w(x). \tag{12.12.26}$$

Solving Equation 12.12.24,

$$w(x) = \frac{L^2 - x^2}{2a^2} + \frac{\kappa L}{ha^2}. \tag{12.12.27}$$

Turning to the transient solution $v(x,t)$, we use separation of variables and find that

$$v(x,t) = \sum_{n=1}^{\infty} C_n \cos\left(\frac{k_n x}{L}\right) \exp\left(-\frac{a^2 k_n^2 t}{L^2}\right), \tag{12.12.28}$$

where k_n is the nth root of the transcendental equation: $k_n \tan(k_n) = hL/\kappa = h^*$. Table 12.12.1 gives the first six roots for various values of hL/κ.

Our final task is to compute C_n. After substituting $t = 0$ into Equation 12.12.28, we are left with a orthogonal expansion of $-w(x)$ using the eigenfunctions $\cos(k_n x/L)$. From Equation 6.3.4,

$$C_n = \frac{\int_0^L -w(x)\cos(k_n x/L)\,dx}{\int_0^L \cos^2(k_n x/L)\,dx} = \frac{-L^3\sin(k_n)/(a^2k_n^3)}{L[k_n+\sin(2k_n)/2]/(2k_n)} = -\frac{2L^2\sin(k_n)}{a^2k_n^2[k_n+\sin(2k_n)/2]}. \tag{12.12.29}$$

Combining Equation 12.12.28 and Equation 12.12.29,

$$v(x,t) = -\frac{2L^2}{a^2}\sum_{n=1}^{\infty}\frac{\sin(k_n)\cos(k_n x/L)}{k_n^2[k_n+\sin(2k_n)/2]}\exp\left(-\frac{a^2k_n^2t}{L^2}\right). \tag{12.12.30}$$

Consequently, $A(x,t)$ equals

$$A(x,t) = \frac{L^2-x^2}{2a^2} + \frac{\kappa L}{ha^2} - \frac{2L^2}{a^2}\sum_{n=1}^{\infty}\frac{\sin(k_n)\cos(k_n x/L)}{k_n^2[k_n+\sin(2k_n)/2]}\exp\left(-\frac{a^2k_n^2t}{L^2}\right). \tag{12.12.31}$$

We now wish to use the solution Equation 12.12.31 to find the temperature distribution within the slab when it is heated by a time-dependent source $f(t)$. As in the case of time-dependent boundary conditions, we imagine that we can break the process into an infinite number of small changes to the heating, which occur at the times $t = \tau_1$, $t = \tau_2$, etc. Consequently, the temperature distribution at the time t following the change at $t = \tau_n$ and before the change at $t = \tau_{n+1}$ is

$$u(x,t) = f(0)A(x,t) + \sum_{k=1}^{n} A(x,t-\tau_k)\frac{\Delta f_k}{\Delta\tau_k}\Delta\tau_k, \tag{12.12.32}$$

where

$$\Delta f_k = f(\tau_k) - f(\tau_{k-1}), \qquad \text{and} \qquad \Delta\tau_k = \tau_k - \tau_{k-1}. \tag{12.12.33}$$

In the limit of $\Delta\tau_k \to 0$,

$$u(x,t) = f(0)A(x,t) + \int_0^t A(x,t-\tau)f'(\tau)\,d\tau = f(t)A(x,0) + \int_0^t f(\tau)\frac{\partial A(x,t-\tau)}{\partial\tau}\,d\tau. \tag{12.12.34}$$

In our present problem,

$$f(t) = \frac{2q}{\rho C_p AL}\sin^2(\omega t), \qquad \text{and} \qquad f'(t) = \frac{2q\omega}{\rho C_p AL}\sin(2\omega t). \tag{12.12.35}$$

Therefore,

$$u(x,t) = \frac{2q\omega}{\rho C_p AL}\int_0^t \sin(2\omega\tau)\left\{\frac{L^2-x^2}{2a^2} + \frac{\kappa L}{ha^2} - \frac{2L^2}{a^2}\sum_{n=1}^{\infty}\frac{\sin(k_n)}{k_n^2[k_n+\sin(2k_n)/2]}\cos\left(\frac{k_n x}{L}\right)\exp\left[-\frac{a^2k_n^2(t-\tau)}{L^2}\right]\right\}d\tau \tag{12.12.36}$$

$$= -\frac{q}{\rho C_p AL}\left(\frac{L^2-x^2}{2a^2} + \frac{\kappa L}{ha^2}\right)\cos(2\omega\tau)\Big|_0^t \tag{12.12.37}$$

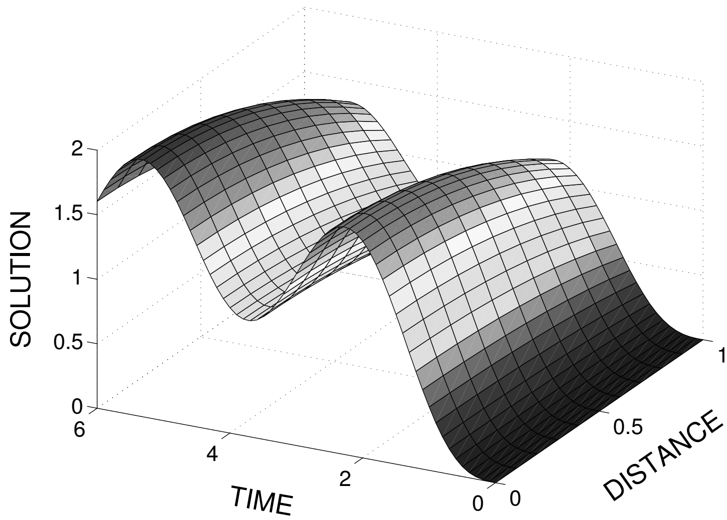


Figure 12.12.1: The nondimensional temperature $a^2 A\rho C_p u(x,t)/qL$ within a slab that we heat by alternating electric current as a function of position x/L and time a^2t/L^2 when we insulate the $x=0$ end and let the $x=L$ end radiate to free space at temperature zero. The initial temperature is zero, $hL/\kappa = 1$, and $a^2/(L^2\omega) = 1$.

$$
\begin{aligned}
&- \frac{4L^2 q\omega}{a^2\rho C_p AL} \sum_{n=1}^{\infty} \frac{\sin(k_n)\exp(-a^2k_n^2t/L^2)}{k_n^2[k_n+\sin(2k_n)/2]} \cos\left(\frac{k_n x}{L}\right) \int_0^t \sin(2\omega\tau)\exp\left(\frac{a^2k_n^2\tau}{L^2}\right)\,d\tau \\
&= \frac{qL}{a^2A\rho C_p}\left\{\left[\frac{L^2-x^2}{2L^2}+\frac{\kappa}{hL}\right][1-\cos(2\omega t)]\right. \\
&\quad - \sum_{n=1}^{\infty} \frac{4\sin(k_n)\cos(k_n x/L)}{k_n^2[k_n+\sin(2k_n)/2][4+a^4k_n^4/(L^4\omega^2)]} \\
&\qquad \left.\times \left[\frac{a^2k_n^2}{\omega L^2}\sin(2\omega t) - 2\cos(2\omega t) + 2\exp\left(-\frac{a^2k_n^2 t}{L^2}\right)\right]\right\}. \qquad \textbf{(12.12.38)}
\end{aligned}
$$

Figure 12.12.1 illustrates Equation 12.12.38 for $hL/\kappa = 1$, and $a^2/(L^2\omega) = 1$. This figure was created using the MATLAB script

```
clear
asq_over_omegaL2 = 1; h_star = 1; m = 0; M = 10;
dx = 0.1; dt = 0.1;
% create initial guess at k_n
zero = zeros(length(M));
for n = 1:10000
  k1 = 0.1*n; k2 = 0.1*(n+1);
  prod = k1 * tan(k1);
  y1 = h_star - prod; y2 = h_star - k2 * tan(k2);
  if (y1*y2 <= 0 & prod < 2 & m < M) m = m+1; zero(m) = k1; end;
end;
% use Newton-Raphson method to improve values of k_n
for n = 1:M; for k = 1:10
  f = h_star - zero(n) * tan(zero(n));
  fp = - tan(zero(n)) - zero(n) * sec(zero(n))^2;
  zero(n) = zero(n) - f / fp;
```

```
end; end;
% compute grid and initialize solution
X = [0:dx:1]; T = [0:dt:6];
temp1 = (0.5 + 1/h_star)*ones(1,length(X)) - 0.5*X.*X;
temp2 = ones(1,length(T)) - cos(2*T);
u = temp1' * temp2;
XX = X' * ones(1,length(T));
TT = ones(1,length(X))' * T;
% compute solution from Equation 12.12.38
for m = 1:M
  xtemp1 = zero(m) * zero(m);
  xtemp2 = 4 + asq_over_omegaL2*asq_over_omegaL2*xtemp1*xtemp1;
  xtemp3 = asq_over_omegaL2 * xtemp1;
  xtemp4 = zero(m) + sin(2*zero(m))/2;
  xtemp5 = asq_over_omegaL2 * xtemp1;
  aaaaa = 4 * sin(zero(m)) / (xtemp1 * xtemp2 * xtemp4);
  u = u - aaaaa * cos(zero(m)*X)' ...
    * (xtemp5 * sin(2*T) - 2 * cos(2*T) + 2 * exp(-xtemp5 * T));
end
surf(XX,TT,u)
xlabel('DISTANCE','Fontsize',20); ylabel('TIME','Fontsize',20)
zlabel('SOLUTION','Fontsize',20)
```

The oscillating solution, reflecting the periodic heating by the alternating current, rapidly reaches equilibrium. Because heat is radiated to space at $x = L$, the temperature is maximum at $x = 0$ at any given instant as heat flows from $x = 0$ to $x = L$. □

• Example 12.12.2

Consider the following heat conduction problem with time-dependent forcing and/or boundary conditions:

$$\frac{\partial u}{\partial t} = a^2 L(u) + f(P,t), \qquad 0 < t, \tag{12.12.39}$$

$$B(u) = g(Q,t), \qquad 0 < t, \tag{12.12.40}$$

and

$$u(P,0) = h(P), \tag{12.12.41}$$

where

$$L(u) = C_0 + C_1 \frac{\partial}{\partial x_1}\left(K_1 \frac{\partial u}{\partial x_1}\right) + C_2 \frac{\partial}{\partial x_2}\left(K_2 \frac{\partial u}{\partial x_2}\right) + C_3 \frac{\partial}{\partial x_3}\left(K_3 \frac{\partial u}{\partial x_3}\right), \tag{12.12.42}$$

$$B(u) = c_0 + c_1 \frac{\partial u}{\partial x_1} + c_2 \frac{\partial u}{\partial x_2} + c_3 \frac{\partial u}{\partial x_3}, \tag{12.12.43}$$

P denotes an arbitrary interior point at (x_1, x_2, x_3) of a region R, and Q is any point on the boundary of R. Here c_i, C_i, and K_i are functions of x_1, x_2, and x_3 only.

Many years ago, Bartels and Churchill[63] extended Duhumel's theorem to solve this heat conduction problem. They did this by first introducing the simpler initial-boundary-value

[63] Bartels, R. C. F., and R. V. Churchill, 1942: Resolution of boundary problems by the use of a generalized convolution. *Am. Math. Soc. Bull.*, **48**, 276–282.

problem:

$$\frac{\partial v}{\partial t} = a^2 L(v) + f(P, t_1), \qquad 0 < t, \tag{12.12.44}$$

$$B(v) = g(Q, t_1), \qquad 0 < t, \tag{12.12.45}$$

and

$$v(P, 0) = h(P), \tag{12.12.46}$$

which has a constant forcing and boundary conditions in place of the time-dependent ones. Here t_1 denotes an arbitrary but *fixed* instant of time. Then Bartels and Churchill proved that the solution to the original problem is given by the convolution integral

$$u(P, t) = \frac{\partial}{\partial t}\left[\int_0^t v(P, t - \tau, \tau)\, d\tau\right]. \tag{12.12.47}$$

To illustrate[64] this technique, let us solve

$$\frac{\partial u}{\partial t} = a^2\left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r}\right) = \frac{a^2}{r}\frac{\partial^2 (ru)}{\partial r^2}, \qquad \alpha < r < \beta, \quad 0 < t, \tag{12.12.48}$$

subject to the boundary conditions

$$u(\alpha, t) = u_0 e^{-ct}, \quad \frac{\partial u(\beta, t)}{\partial r} = 0, \qquad 0 < t, \tag{12.12.49}$$

and the initial condition $u(r, 0) = u_0$, $\alpha < r < \beta$.

We begin by solving the alternative problem

$$\frac{\partial v}{\partial t} = a^2\left(\frac{\partial^2 v}{\partial r^2} + \frac{2}{r}\frac{\partial v}{\partial r}\right) = \frac{a^2}{r}\frac{\partial^2 (rv)}{\partial r^2}, \qquad \alpha < r < \beta, \quad 0 < t, \tag{12.12.50}$$

subject to the boundary conditions

$$v(\alpha, t, t') = u_0 e^{-ct'}, \quad \frac{\partial v(\beta, t, t')}{\partial r} = 0, \qquad 0 < t, \tag{12.12.51}$$

and the initial condition $v(r, 0, t') = u_0$, $\alpha < r < \beta$, or equivalently

$$\frac{\partial w}{\partial t} = a^2\left(\frac{\partial^2 w}{\partial r^2} + \frac{2}{r}\frac{\partial w}{\partial r}\right) = \frac{a^2}{r}\frac{\partial^2 (rw)}{\partial r^2}, \qquad \alpha < r < \beta, \quad 0 < t, \tag{12.12.52}$$

subject to the boundary conditions

$$w(\alpha, t, t') = 0, \quad \frac{\partial w(\beta, t, t')}{\partial r} = 0, \qquad 0 < t, \tag{12.12.53}$$

and the initial condition $w(r, 0, t') = u_0(1 - e^{-ct'})$, $\alpha < r < \beta$, where $v(r, t, t') = u_0 e^{-ct'} + w(r, t, t')$.

---

[64] See Reiss, H., and V. K. LaMer, 1950: Diffusional boundary value problems involving moving boundaries, connected with the growth of colloidal particles. *J. Chem. Phys.*, **18**, 1–12.

The heat condition problem Equation 12.12.52 and Equation 12.12.53 can be solved using separation of variables. Following Example 8.3.6, we find that

$$w(r,t,t') = \frac{\alpha u_0(1-e^{-ct'})}{r} \sum_{n=1}^{\infty} \frac{\sin[k_n(r-\alpha)]}{k_n c_n} e^{-a^2 k_n^2 t}, \qquad \mathbf{(12.12.54)}$$

where k_n is the nth root of $\beta k = \tan[k(\beta-\alpha)]$, and $2c_n = \{\beta \sin^2[k_n(\beta-\alpha)] - \alpha\}$. Therefore,

$$u(x,t) = \frac{\alpha u_0}{r} \frac{\partial}{\partial t} \left\{ \int_0^t (1-e^{-c\tau}) \sum_{n=1}^{\infty} \frac{\sin[k_n(r-\alpha)]}{k_n c_n} e^{-a^2 k_n^2 (t-\tau)}\, d\tau \right\} \qquad \mathbf{(12.12.55)}$$

$$= \frac{\alpha u_0}{r} \sum_{n=1}^{\infty} \frac{\sin[k_n(r-\alpha)]}{k_n c_n} \frac{\partial}{\partial t} \left\{ \int_0^t e^{-a^2 k_n^2 (t-\tau)} - e^{-a^2 k_n^2 (t-\tau) - c\tau}\, d\tau \right\} \qquad \mathbf{(12.12.56)}$$

$$= \frac{\alpha u_0}{r} \sum_{n=1}^{\infty} \frac{\sin[k_n(r-\alpha)]}{k_n c_n} \frac{\partial}{\partial t} \left\{ \frac{1-e^{-a^2 k_n^2 t}}{a^2 k_n^2} - \frac{e^{-ct} - e^{-a^2 k_n^2 t}}{a^2 k_n^2 - c} \right\} \qquad \mathbf{(12.12.57)}$$

$$= \frac{\alpha c u_0}{r} \sum_{n=1}^{\infty} \frac{\sin[k_n(r-\alpha)]}{k_n c_n} \, \frac{e^{-a^2 k_n^2 t} - e^{-ct}}{c - a^2 k_n^2}, \qquad \mathbf{(12.12.58)}$$

and the final answer is

$$u(x,t) = u_0 e^{-ct} + \frac{\alpha c u_0}{r} \sum_{n=1}^{\infty} \frac{e^{-a^2 k_n^2 t} - e^{-ct}}{(c - a^2 k_n^2) k_n c_n} \sin[k_n(r-\alpha)]. \qquad \mathbf{(12.12.59)}$$

Problems

1. Solve the heat equation[65]

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

subject to the boundary conditions $u(0,t) = u(L,t) = f(t)$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < L$.

Step 1: First solve the heat conduction problem

$$\frac{\partial A}{\partial t} = a^2 \frac{\partial^2 A}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t,$$

subject to the boundary conditions $A(0,t) = A(L,t) = 1$, $0 < t$, and the initial condition $A(x,0) = 0$, $0 < x < L$. Show that

$$A(x,t) = 1 - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sin[(2n-1)\pi x/L]}{2n-1} e^{-a^2(2n-1)^2\pi^2 t/L^2}.$$

[65] See Tao, L. N., 1960: Magnetohydrodynamic effects on the formation of Couette flow. *J. Aerosp. Sci.*, **27**, 334–338.

Step 2: Use Duhamel's theorem and show that

$$u(x,t) = \frac{4\pi a^2}{L^2} \sum_{n=1}^{\infty} (2n-1) \sin\left[\frac{(2n-1)\pi x}{L}\right] e^{-a^2(2n-1)^2\pi^2 t/L^2} \int_0^t f(\tau) e^{a^2(2n-1)^2\pi^2\tau/L^2}\, d\tau.$$

2. A thermometer measures temperature by the thermal expansion of a liquid (usually mercury or alcohol) stored in a bulb into a glass stem containing an empty cylindrical channel. Under normal conditions, temperature changes occur sufficiently slowly so that the temperature within the liquid is uniform. However, for rapid temperature changes (such as those that would occur during the rapid ascension of an airplane or meteorological balloon), significant errors could occur. In such situations the recorded temperature would lag behind the actual temperature because of the time needed for the heat to conduct in or out of the bulb. During his investigation of this question, McLeod[66] solved

$$\frac{\partial u}{\partial t} = a^2 \frac{1}{r} \frac{\partial}{\partial r}\left(r \frac{\partial u}{\partial r}\right), \qquad 0 \le r < b, \quad 0 < t,$$

subject to the boundary conditions $\lim_{r\to 0} |u(r,t)| < \infty$, and $u(b,t) = \varphi(t)$, $0 < t$, and the initial condition $u(r,0) = 0$, $0 \le r < b$. The analysis was as follows:

Step 1: First solve the heat conduction problem

$$\frac{\partial A}{\partial t} = \frac{a^2}{r} \frac{\partial}{\partial r}\left(r \frac{\partial A}{\partial r}\right), \qquad 0 \le r < b, \quad 0 < t,$$

subject to the boundary conditions $\lim_{r\to 0} |A(r,t)| < \infty$, and $A(b,t) = 1$, $0 < t$, and the initial condition $A(r,0) = 0$, $0 \le r < b$. Show that

$$A(r,t) = 1 - 2 \sum_{n=1}^{\infty} \frac{J_0(k_n r/b)}{k_n\, J_1(k_n)} e^{-a^2 k_n^2 t/b^2},$$

where $J_0(k_n) = 0$.

Step 2: Use Duhamel's theorem and show that

$$u(r,t) = \frac{2a^2}{b^2} \sum_{n=1}^{\infty} \frac{k_n\, J_0(k_n r/b)}{J_1(k_n)} \int_0^t \varphi(\tau)\, e^{-a^2 k_n^2 (t-\tau)/b^2}\, d\tau.$$

Step 3: If $\varphi(t) = Gt$, show that

$$u(r,t) = 2G \sum_{n=1}^{\infty} \frac{J_0(k_n r/b)}{k_n\, J_1(k_n)} \left[t + \frac{b^2}{a^2 k_n^2} \left(e^{-a^2 k_n^2 t/b^2} - 1 \right) \right].$$

McLeod found that for a mercury thermometer of 10-cm length a lag of 0.01°C would occur for a warming rate of 0.032°C s^{-1} (a warming gradient of 1.9°C per thousand feet

[66] McLeod, A. R., 1919: On the lags of thermometers with spherical and cylindrical bulbs in a medium whose temperature is changing at a constant rate. *Philos. Mag., Ser. 6*, **37**, 134–144. See also Bromwich, T. J. I'A., 1919: Examples of operational methods in mathematical physics. *Philos. Mag., Ser. 6*, **37**, 407–419; McLeod, A. R., 1922: On the lags of thermometers. *Philos. Mag., Ser. 6*, **43**, 49–70.

and a descent of one thousand feet per minute). Although this is a very small number, when he included the surface conductance of the glass tube, the lag increased to 0.85°C. Similar problems plague bimetal thermometers[67] and thermistors[68] used in radiosondes (meteorological sounding balloons).

3. A classic problem[69] in fluid mechanics is the motion of a semi-infinite viscous fluid that results from the sudden movement of the adjacent wall starting at $t = 0$. Initially the fluid is at rest. If we denote the velocity of the fluid parallel to the wall by $u(x,t)$, the governing equation is

$$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

with the boundary conditions $u(0,t) = V(t)$, $\lim_{x\to\infty} u(x,t) \to 0$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < \infty$.

Step 1: Find the step response by solving

$$\frac{\partial A}{\partial t} = \nu \frac{\partial^2 A}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions

$$A(0,t) = 1, \quad \lim_{x\to\infty} A(x,t) \to 0, \qquad 0 < t,$$

and the initial condition $A(x,0) = 0$, $0 < x < \infty$. Show that

$$A(x,t) = \operatorname{erfc}\left(\frac{x}{2\sqrt{\nu t}}\right) = \frac{2}{\sqrt{\pi}} \int_{x/2\sqrt{\nu t}}^{\infty} e^{-\eta^2}\, d\eta,$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function. Hint: Use Laplace transforms.

Step 2: Use Duhamel's theorem and show that the solution is

$$u(x,t) = \int_0^t V(t-\tau) \frac{x \exp(-x^2/4\nu\tau)}{2\sqrt{\pi\nu\tau^3}}\, d\tau = \frac{2}{\pi} \int_{x/\sqrt{4\nu t}}^{\infty} V\left(t - \frac{x^2}{4\nu\eta^2}\right) e^{-\eta^2}\, d\eta.$$

4. During their study of the propagation of a temperature step in a nearly supercritical van der Waals gas, Zappoli and Durand-Daubin[70] solved

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

with the boundary conditions $u(0,t) = u_0 - \frac{2}{3} f(t)$, $\lim_{x\to\infty} u(x,t) \to 0$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < \infty$, where u_0 is a constant.

---

[67] Mitra, H., and M. B. Datta, 1954: Lag coefficient of bimetal thermometer of chronometric radiosonde. *Indian J. Meteorol. Geophys.*, **5**, 257–261.

[68] Badgley, F. I., 1957: Response of radiosonde thermistors. *Rev. Sci. Instrum.*, **28**, 1079–1084.

[69] This problem was first posed and partially solved by Stokes, G. G., 1850: On the effect of the internal friction of fluids on the motions of pendulums. *Proc. Cambridge Philos. Soc.*, **9, Part II**, [8]–[106].

[70] Zappoli, B., and A. Durand-Daubin, 1994: Heat and mass transport in a near supercritical fluid. *Phys. Fluids*, **6**, 1929–1936.

Step 1: Find the step response by solving

$$\frac{\partial A}{\partial t} = a^2 \frac{\partial^2 A}{\partial x^2}, \qquad 0 < x < \infty, \quad 0 < t,$$

subject to the boundary conditions $A(0,t) = 1$, $\lim_{x\to\infty} A(x,t) \to 0$, $0 < t$, and the initial condition $A(x,0) = 0$, $0 < x < \infty$. Show that

$$A(x,t) = \operatorname{erfc}\left(\frac{x}{2a\sqrt{t}}\right) = \frac{2}{\sqrt{\pi}} \int_{x/(2a\sqrt{t}\,)}^{\infty} e^{-\eta^2}\, d\eta,$$

where erfc(·) is the complementary error function. Hint: Use Laplace transforms.

Step 2: Use Duhamel's theorem and show that the solution is

$$u(x,t) = u_0 \operatorname{erfc}\left(\frac{x}{2a\sqrt{t}}\right) - \frac{4}{3\sqrt{\pi}} \int_{x/(2a\sqrt{t}\,)}^{\infty} f\left(t - \frac{x^2}{4a^2\eta^2}\right) e^{-\eta^2}\, d\eta.$$

5. Solve the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions $u(0,t) = f(t)$, $u_x(1,t) = -hu(1,t)$, $0 < t$, and the initial condition $u(x,0) = 0$, $0 < x < 1$.

Step 1: First solve the heat conduction problem

$$\frac{\partial A}{\partial t} = \frac{\partial^2 A}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t,$$

subject to the boundary conditions $A(0,t) = 1$, $A_x(1,t) = -hA(1,t)$, $0 < t$, and the initial condition $A(x,0) = 0$, $0 < x < 1$. Show that

$$A(x,t) = 1 - \frac{hx}{1+h} - 2\sum_{n=1}^{\infty} \frac{k_n^2 + h^2}{k_n\,(k_n^2 + h^2 + h)} \sin(k_n x) e^{-k_n^2 t},$$

where k_n is the nth root of $k\cot(k) = -h$.

Step 2: Use Duhamel's theorem and show that

$$u(x,t) = 2\sum_{n=1}^{\infty} \frac{k_n(k_n^2 + h^2)}{k_n^2 + h^2 + h} \sin(k_n x) e^{-k_n^2 t} \int_0^t f(\tau) e^{k_n^2 \tau}\, d\tau.$$

12.13 THE SOLUTION OF LAPLACE'S EQUATION

Laplace transforms are useful in solving Laplace's or Poisson's equation over a semi-infinite strip. The following problem illustrates this technique.

Let us solve Poisson's equation within a semi-infinite circular cylinder

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial z^2} = \frac{2}{b} n(z)\delta(r-b), \quad 0 \le r < a, \quad 0 < z < \infty, \qquad \mathbf{(12.13.1)}$$

subject to the boundary conditions

$$u(r,0) = 0, \quad \lim_{z\to\infty} |u(r,z)| < \infty, \quad 0 \le r < a, \tag{12.13.2}$$

and

$$u(a,z) = 0, \qquad 0 < z < \infty, \tag{12.13.3}$$

where $0 < b < a$. This problem gives the electrostatic potential within a semi-infinite cylinder of radius a that is grounded and has the charge density of $n(z)$ within an infinitesimally thin shell located at $r = b$.

Because the domain is semi-infinite in the z direction, we introduce the Laplace transform

$$U(r,s) = \int_0^\infty u(r,z)\, e^{-sz}\, dz. \tag{12.13.4}$$

Thus, taking the Laplace transform of Equation 12.13.1, we have that

$$\frac{1}{r}\frac{d}{dr}\left[r\frac{dU(r,s)}{dr}\right] + s^2U(r,s) - su(r,0) - u_z(r,0) = \frac{2}{b}N(s)\delta(r-b). \tag{12.13.5}$$

Although $u(r,0) = 0$, $u_z(r,0)$ is unknown and we denote its value by $f(r)$. Therefore, Equation 12.13.5 becomes

$$\frac{1}{r}\frac{d}{dr}\left[r\frac{dU(r,s)}{dr}\right] + s^2U(r,s) = f(r) + \frac{2}{b}N(s)\delta(r-b), \quad 0 \le r < a, \tag{12.13.6}$$

with $\lim_{r\to 0} |U(r,s)| < \infty$, and $U(a,s) = 0$.

To solve Equation 12.13.6 we first assume that we can rewrite $f(r)$ as the Fourier-Bessel series

$$f(r) = \sum_{n=1}^\infty A_n J_0(k_n r/a), \tag{12.13.7}$$

where k_n is the nth root of the $J_0(k) = 0$, and

$$A_n = \frac{2}{a^2 J_1^2(k_n)} \int_0^a f(r)\, J_0(k_n r/a)\, r\, dr. \tag{12.13.8}$$

Similarly, the expansion for the delta function is

$$\delta(r-b) = \frac{2b}{a^2} \sum_{n=1}^\infty \frac{J_0(k_n b/a) J_0(k_n r/a)}{J_1^2(k_n)}, \tag{12.13.9}$$

because

$$\int_0^a \delta(r-b) J_0(k_n r/a)\, r\, dr = b\, J_0(k_n b/a). \tag{12.13.10}$$

Why we chose this particular expansion will become apparent shortly.

Thus, Equation 12.13.6 may be rewritten as

$$\frac{1}{r}\frac{d}{dr}\left[r\frac{dU(r,s)}{dr}\right] + s^2U(r,s) = \frac{2}{a^2} \sum_{n=1}^\infty \frac{2N(s)J_0(k_n b/a) + a_k}{J_1^2(k_n)} J_0(k_n r/a), \tag{12.13.11}$$

where $a_k = \int_0^a f(r)\, J_0(k_n r/a)\, r\, dr$.

The form of the right side of Equation 12.13.11 suggests that we seek solutions of the form

$$U(r,s)=\sum_{n=1}^{\infty} B_n J_0(k_n r/a), \qquad 0 \le r < a. \tag{12.13.12}$$

We now understand why we rewrote the right side of Equation 12.13.6 as a Fourier-Bessel series; the solution $U(r,s)$ automatically satisfies the boundary condition $U(a,s)=0$. Substituting Equation 12.13.12 into Equation 12.13.11, we find that

$$U(r,s)=\frac{2}{a^2}\sum_{n=1}^{\infty}\frac{2N(s)J_0(k_n b/a)+a_k}{(s^2-k_n^2/a^2)J_1^2(k_n)}J_0(k_n r/a), \qquad 0 \le r < a. \tag{12.13.13}$$

We have not yet determined a_k. Note, however, that in order for the inverse of Equation 12.13.13 *not* to grow as $e^{k_n z/a}$, the numerator must vanish when $s=k_n/a$ and $s=k_n/a$ is a removable pole. Thus,

$$a_k=-2N(k_n/a)J_0(k_n b/a), \tag{12.13.14}$$

and

$$U(r,s)=\frac{4}{a^2}\sum_{n=1}^{\infty}\frac{[N(s)-N(k_n/a)]J_0(k_n b/a)}{(s^2-k_n^2/a^2)J_1^2(k_n)}J_0(k_n r/a), \qquad 0 \le r < a. \tag{12.13.15}$$

The inverse of $U(r,s)$ then follows directly from simple inversions, the convolution theorem, and the definition of the Laplace transform. The complete solution is

$$\begin{aligned}u(r,z)=&\frac{2}{a}\sum_{n=1}^{\infty}\frac{J_0(k_n b/a)J_0(k_n r/a)}{k_n\, J_1^2(k_n)}\\ &\times\left[\int_0^z n(\tau)e^{k_n(z-\tau)/a}\,d\tau-\int_0^z n(\tau)e^{-k_n(z-\tau)/a}\,d\tau\right.\\ &\qquad\left.-\int_0^{\infty} n(\tau)e^{-k_n\tau/a}e^{k_n z/a}\,d\tau+\int_0^{\infty} n(\tau)e^{-k_n\tau/a}e^{-k_n z/a}\,d\tau\right] \qquad (12.13.16)\\ =&\frac{2}{a}\sum_{n=1}^{\infty}\frac{J_0(k_n b/a)J_0(k_n r/a)}{k_n\, J_1^2(k_n)} \qquad (12.13.17)\\ &\times\left[\int_0^{\infty} n(\tau)e^{-k_n(z+\tau)/a}\,d\tau-\int_0^z n(\tau)e^{-k_n(z-\tau)/a}\,d\tau-\int_z^{\infty} n(\tau)e^{-k_n(\tau-z)/a}\,d\tau\right].\end{aligned}$$

Problems

1. Use Laplace transforms to solve

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0, \qquad 0<x<\infty, \quad 0<y<a,$$

subject to the boundary conditions

$$u(0,y)=1, \quad \lim_{x\to\infty}|u(x,y)|<\infty, \qquad 0<y<a,$$

and

$$u(x,0) = u(x,a) = 0, \qquad 0 < x < \infty.$$

2. Use Laplace transforms to solve

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial z^2} = 0, \qquad 0 \le r < a, \quad 0 < z < \infty,$$

subject to the boundary conditions

$$u(r,0) = 1, \quad \lim_{z\to\infty} |u(r,z)| < \infty, \qquad 0 \le r < a,$$

and

$$\lim_{r\to 0} |u(r,z)| < \infty, \quad u(a,z) = 0, \qquad 0 < z < \infty.$$

Further Readings

Churchill, R. V., 1972: *Operational Mathematics*. McGraw-Hill Book Co., 481 pp. A classic textbook on Laplace transforms.

Doetsch, G., 1950: *Handbuch der Laplace-Transformation. Band 1. Theorie der Laplace-Transformation*. Birkhäuser Verlag, 581 pp.; Doetsch, G., 1955: *Handbuch der Laplace-Transformation. Band 2. Anwendungen der Laplace-Transformation. 1. Abteilung*. Birkhäuser Verlag, 433 pp.; Doetsch, G., 1956: *Handbuch der Laplace-Transformation. Band 3. Anwendungen der Laplace-Transformation. 2. Abteilung*. Birkhäuser Verlag, 298 pp. One of the standard reference books on Laplace transforms.

LePage, W. R., 1980: *Complex Variables and the Laplace Transform for Engineers.* Dover, 483 pp. Laplace transforms approached from complex variable theory and a few applications.

McLachlan, N. W., 1944: *Complex Variables & Operational Calculus with Technical Applications.* Cambridge Press, 355 pp. An early book on Laplace transforms from the view of complex variables.

Thomson, W. T., 1960: *Laplace Transformation.* Prentice-Hall, 255 pp. Presents Laplace transforms in the engineering applications that gave them birth.

Watson, E. J., 1981: *Laplace Transforms and Applications*. Van Nostrand Reinhold, 205 pp. A brief and very complete guide to Laplace transforms.

Chapter 13
The Z-Transform

Since the Second World War, the rise of digital technology has resulted in a corresponding demand for designing and understanding discrete-time (data sampled) systems. These systems are governed by *difference equations* in which members of the sequence y_n are coupled to each other.

One source of difference equations is the numerical evaluation of integrals on a digital computer. Because we can only have values at discrete time points $t_k = kT$ for $k = 0, 1, 2, \ldots$, the value of the integral $y(t) = \int_0^t f(\tau)\, d\tau$ is

$$y(kT) = \int_0^{kT} f(\tau)\, d\tau = \int_0^{(k-1)T} f(\tau)\, d\tau + \int_{(k-1)T}^{kT} f(\tau)\, d\tau \qquad (\mathbf{13.0.1})$$

$$= y[(k-1)T] + \int_{(k-1)T}^{kT} f(\tau)\, d\tau = y[(k-1)T] + Tf(kT), \qquad (\mathbf{13.0.2})$$

because $\int_{(k-1)T}^{kT} f(\tau)\, d\tau \approx Tf(kT)$. The right side of Equation 13.0.2 is an example of a first-order difference equation because the numerical scheme couples the sequence value $y(kT)$ directly to the previous sequence value $y[(k-1)T]$. If Equation 13.0.2 had contained $y[(k-2)T]$, then it would have been a second-order difference equation, and so forth.

Although we could use the conventional Laplace transform to solve these difference equations, the use of z-transforms can greatly facilitate the analysis, especially when we only desire responses at the sampling instants. Often the entire analysis can be done using only the transforms and the analyst does not actually find the sequence $y(kT)$.

In this chapter we will first define the z-transform and discuss its properties. Then we will show how to find its inverse. Finally we shall use them to solve difference equations.

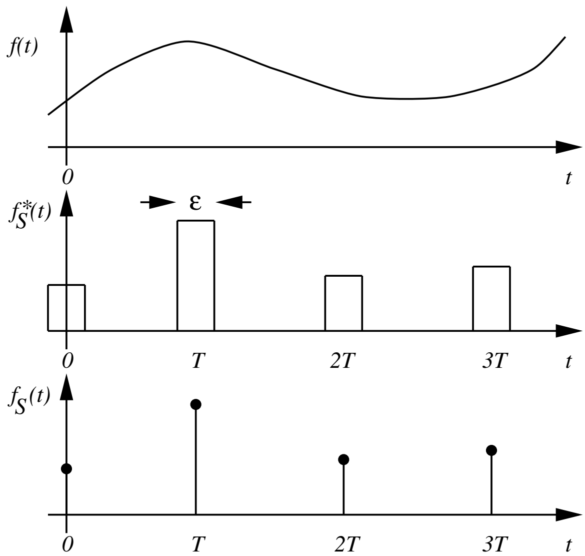


Figure 13.1.1: Schematic of how a continuous function $f(t)$ is sampled by a narrow-width pulse sampler $f_S^*(t)$ and an ideal sampler $f_S(t)$.

13.1 THE RELATIONSHIP OF THE Z-TRANSFORM TO THE LAPLACE TRANSFORM[1]

Let $f(t)$ be a continuous function that an instrument samples every T units of time. We denote this data-sampled function by $f_S^*(t)$. See Figure 13.1.1. Taking ϵ, the duration of an individual sampling event, to be small, we may approximate the narrow-width pulse in Figure 13.1.1 by flat-topped pulses. Then $f_S^*(t)$ approximately equals

$$f_S^*(t) \approx \frac{1}{\epsilon} \sum_{n=0}^{\infty} f(nT) \left[H(t - nT + \epsilon/2) - H(t - nT - \epsilon/2)\right], \tag{13.1.1}$$

if $\epsilon \ll T$.

Clearly the presence of ϵ is troublesome in Equation 13.1.1; it adds one more parameter to our problem. For this reason we introduce the concept of the *ideal sampler*, where the sampling time becomes infinitesimally small so that

$$f_S(t) = \lim_{\epsilon \to 0} \sum_{n=0}^{\infty} f(nT) \left[\frac{H(t - nT + \epsilon/2) - H(t - nT - \epsilon/2)}{\epsilon}\right] = \sum_{n=0}^{\infty} f(nT)\delta(t - nT). \tag{13.1.2}$$

Let us now find the Laplace transform of this data-sampled function. From the linearity property of Laplace transforms,

$$F_S(s) = \mathcal{L}[f_S(t)] = \mathcal{L}\left[\sum_{n=0}^{\infty} f(nT)\delta(t - nT)\right] = \sum_{n=0}^{\infty} f(nT)\mathcal{L}[\delta(t - nT)]. \tag{13.1.3}$$

---

[1] Gera (Gera, A. E., 1999: The relationship between the z-transform and the discrete-time Fourier transform. *IEEE Trans. Auto. Control*, **AC-44**, 370–371) explored the general relationship between the one-sided discrete-time Fourier transform and the one-sided z-transform. See also Naumović, M. B., 2001: Interrelationship between the one-sided discrete-time Fourier transform and one-sided delta transform. *Electr. Engng.*, **83**, 99–101.

Because $\mathcal{L}[\delta(t - nT)] = e^{-nsT}$, Equation 13.1.3 simplifies to

$$F_S(s) = \sum_{n=0}^{\infty} f(nT)e^{-nsT}. \tag{13.1.4}$$

If we now make the substitution that $z = e^{sT}$, then $F_S(s)$ becomes

$$\boxed{F(z) = \mathcal{Z}(f_n) = \sum_{n=0}^{\infty} f_n z^{-n},} \tag{13.1.5}$$

where $F(z)$ is the one-sided z-transform[2] of the sequence $f(nT)$, which we shall now denote by f_n. Here $\mathcal{Z}$ denotes the operation of taking the z-transform while $\mathcal{Z}^{-1}$ represents the inverse z-transformation. We will consider methods for finding the inverse z-transform in Section 13.3.

Just as the Laplace transform was defined by an integration in t, the z-transform is defined by a power series (Laurent series) in z. Consequently, every z-transform has a region of convergence that must be implicitly understood if not explicitly stated. Furthermore, just as the Laplace integral diverged for certain functions, there are sequences where the associated power series diverges and its z-transform does not exist.

Consider now the following examples of how to find the z-transform.

• Example 13.1.1

Given the unit sequence $f_n = 1$, $n \ge 0$, let us find $F(z)$. Substituting f_n into the definition of the z-transform leads to

$$F(z) = \sum_{n=0}^{\infty} z^{-n} = \frac{z}{z-1}, \tag{13.1.6}$$

because $\sum_{n=0}^{\infty} z^{-n}$ is a complex-valued *geometric series* with common ratio z^{-1}. This series converges if $|z^{-1}| < 1$ or $|z| > 1$, which gives the region of convergence of $F(z)$.

MATLAB's symbolic toolbox provides an alternative to the hand computation of the z-transform. In the present case, the command

```
>> syms z; syms n positive
>> ztrans(1,n,z)
```

yields

```
ans =
z/(z-1)
```

□

• Example 13.1.2

Let us find the z-transform of the sequence

$$f_n = e^{-anT}, \qquad n \ge 0, \tag{13.1.7}$$

[2] The standard reference is Jury, E. I., 1964: *Theory and Application of the z-Transform Method*. John Wiley & Sons, 330 pp.

for a real and a imaginary.

For a real, substitution of the sequence into the definition of the z-transform yields

$$F(z) = \sum_{n=0}^{\infty} e^{-anT} z^{-n} = \sum_{n=0}^{\infty} \left(e^{-aT} z^{-1}\right)^n. \qquad (13.1.8)$$

If $u = e^{-aT} z^{-1}$, then Equation 13.1.8 is a geometric series so that

$$F(z) = \sum_{n=0}^{\infty} u^n = \frac{1}{1-u}. \qquad (13.1.9)$$

Because $|u| = e^{-aT}|z^{-1}|$, the condition for convergence is that $|z| > e^{-aT}$. Thus,

$$F(z) = \frac{z}{z - e^{-aT}}, \qquad |z| > e^{-aT}. \qquad (13.1.10)$$

For imaginary a, the infinite series in Equation 13.1.8 converges if $|z| > 1$, because $|u| = |z^{-1}|$ when a is imaginary. Thus,

$$F(z) = \frac{z}{z - e^{-aT}}, \qquad |z| > 1. \qquad (13.1.11)$$

Although the z-transforms in Equation 13.1.10 and Equation 13.1.11 are the same in these two cases, the corresponding regions of convergence are different. If a is a complex number, then

$$F(z) = \frac{z}{z - e^{-aT}}, \qquad |z| > |e^{-aT}|. \qquad (13.1.12)$$

Checking our work using MATLAB, we type the commands:

```
>> syms a z; syms n T positive
>> ztrans(exp(-a*n*T),n,z);
>> simplify(ans)
```

which yields

```
ans =
z*exp(a*T)/(z*exp(a*T)-1)
```

□

• Example 13.1.3

Let us find the z-transform of the sinusoidal sequence

$$f_n = \cos(n\omega T), \qquad n \ge 0. \qquad (13.1.13)$$

Substituting Equation 13.1.13 into the definition of the z-transform results in

$$F(z) = \sum_{n=0}^{\infty} \cos(n\omega T) z^{-n}. \qquad (13.1.14)$$

From Euler's formula,

$$\cos(n\omega T) = \tfrac{1}{2}(e^{in\omega T} + e^{-in\omega T}), \qquad (13.1.15)$$

so that Equation 13.1.14 becomes

$$F(z) = \frac{1}{2}\sum_{n=0}^{\infty}\left(e^{in\omega T}z^{-n} + e^{-in\omega T}z^{-n}\right), \tag{13.1.16}$$

or

$$F(z) = \tfrac{1}{2}\left[\mathcal{Z}(e^{in\omega T}) + \mathcal{Z}(e^{-in\omega T})\right]. \tag{13.1.17}$$

From Equation 13.1.11,

$$\mathcal{Z}(e^{\pm in\omega T}) = \frac{z}{z - e^{\pm i\omega T}}, \qquad |z| > 1. \tag{13.1.18}$$

Substituting Equation 13.1.18 into Equation 13.1.17 and simplifying yields

$$F(z) = \frac{z[z - \cos(\omega T)]}{z^2 - 2z\cos(\omega T) + 1}, \qquad |z| > 1. \tag{13.1.19}$$

□

• **Example 13.1.4**

Let us find the z-transform for the sequence

$$f_n = \begin{cases} 1, & 0 \le n \le 5, \\ (\frac{1}{2})^n, & 6 \le n. \end{cases} \tag{13.1.20}$$

From the definition of the z-transform,

$$\mathcal{Z}(f_n) = F(z) = \sum_{n=0}^{5} z^{-n} + \sum_{n=6}^{\infty}\left(\frac{1}{2z}\right)^n. \tag{13.1.21}$$

$$= 1 + \frac{1}{z} + \frac{1}{z^2} + \frac{1}{z^3} + \frac{1}{z^4} + \frac{1}{z^5} + \frac{2z}{2z-1}$$
$$- 1 - \frac{1}{2z} - \frac{1}{4z^2} - \frac{1}{8z^3} - \frac{1}{16z^4} - \frac{1}{32z^5} \tag{13.1.22}$$
$$= \frac{2z}{2z-1} + \frac{1}{2z} + \frac{3}{4z^2} + \frac{7}{8z^3} + \frac{15}{16z^4} + \frac{31}{32z^5}. \tag{13.1.23}$$

We could also have obtained Equation 13.1.23 via MATLAB by typing the commands:

```
>> syms z; syms n positive
>> ztrans('1+((1/2)^n-1)*Heaviside(n-6)',n,z)
```

which yields

```
ans =
2*z/(2*z-1)+1/2/z+3/4/z^2+7/8/z^3+15/16/z^4+31/32/z^5
```

□

We summarize some of the more commonly encountered sequences and their transforms in Table 13.1.1 along with their regions of convergence.

Table 13.1.1: Z-Transforms of Some Commonly Used Sequences

	$\mathbf{f_n},\ \mathbf{n \geq 0}$	$\mathbf{F(z)}$	**Region of convergence**
1.	$f_0 = k = \text{const.}$ $f_n = 0,\ n \geq 1$	k	$\lvert z\rvert > 0$
2.	$f_m = k = \text{const.}$ $f_n = 0$, for all values of $n \neq m$	kz^{-m}	$\lvert z\rvert > 0$
3.	$k = \text{constant}$	$kz/(z-1)$	$\lvert z\rvert > 1$
4.	kn	$kz/(z-1)^2$	$\lvert z\rvert > 1$
5.	kn^2	$kz(z+1)/(z-1)^3$	$\lvert z\rvert > 1$
6.	ke^{-anT}, a complex	$kz/\left(z - e^{-aT}\right)$	$\lvert z\rvert > \lvert e^{-aT}\rvert$
7.	kne^{-anT}, a complex	$\frac{kze^{-aT}}{(z-e^{-aT})^2}$	$\lvert z\rvert > \lvert e^{-aT}\rvert$
8.	$\sin(\omega_0 nT)$	$\frac{z\sin(\omega_0 T)}{z^2-2z\cos(\omega_0 T)+1}$	$\lvert z\rvert > 1$
9.	$\cos(\omega_0 nT)$	$\frac{z[z-\cos(\omega_0 T)]}{z^2-2z\cos(\omega_0 T)+1}$	$\lvert z\rvert > 1$
10.	$e^{-anT}\sin(\omega_0 nT)$	$\frac{ze^{-aT}\sin(\omega_0 T)}{z^2-2ze^{-aT}\cos(\omega_0 T)+e^{-2aT}}$	$\lvert z\rvert > e^{-aT}$
11.	$e^{-anT}\cos(\omega_0 nT)$	$\frac{ze^{-aT}[ze^{aT}-\cos(\omega_0 T)]}{z^2-2ze^{-aT}\cos(\omega_0 T)+e^{-2aT}}$	$\lvert z\rvert > e^{-aT}$
12.	α^n , α constant	$z/(z-\alpha)$	$\lvert z\rvert > \lvert\alpha\rvert$
13.	$n\alpha^n$	$\alpha z/(z-\alpha)^2$	$\lvert z\rvert > \lvert\alpha\rvert$
14.	$n^2\alpha^n$	$\alpha z(z+\alpha)/(z-\alpha)^3$	$\lvert z\rvert > \lvert\alpha\rvert$
15.	$\sinh(\omega_0 nT)$	$\frac{z\sinh(\omega_0 T)}{z^2-2z\cosh(\omega_0 T)+1}$	$\lvert z\rvert > \cosh(\omega_0 T)$
16.	$\cosh(\omega_0 nT)$	$\frac{z[z-\cosh(\omega_0 T)]}{z^2-2z\cosh(\omega_0 T)+1}$	$\lvert z\rvert > \sinh(\omega_0 T)$
17.	$a^n/n!$	$e^{a/z}$	$\lvert z\rvert > 0$
18.	$[\ln(a)]^n/n!$	$a^{1/z}$	$\lvert z\rvert > 0$

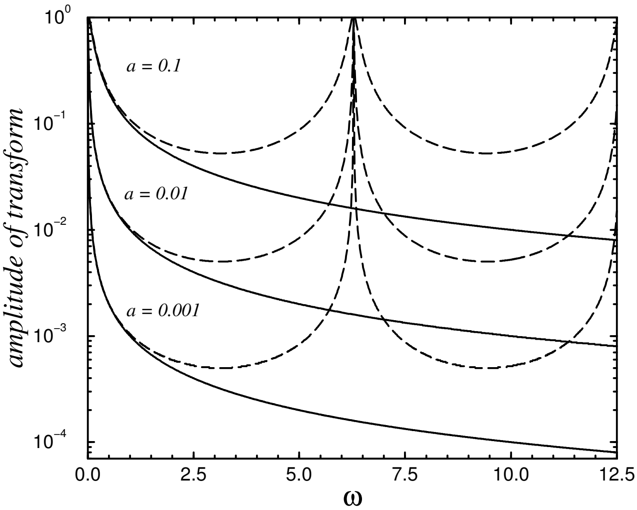


Figure 13.1.2: The amplitude of the Laplace or Fourier transform (solid line) for the function $ae^{-at}H(t)$ and the z-transform (dashed line) for the sequence $f_n = ae^{-anT}$ as a function of frequency ω for various positive values of a and $T = 1$.

• Example 13.1.5

In many engineering studies, the analysis is done entirely using transforms without actually finding any inverses. Consequently, it is useful to compare and contrast how various transforms behave in very simple test problems.

Consider the time function $f(t) = ae^{-at}H(t)$, $a > 0$. Its Laplace and Fourier transform are identical, namely $a/(a+i\omega)$, if we set $s = i\omega$. In Figure 13.1.2 we illustrate its behavior as a function of positive ω.

Let us now generate the sequence of observations that we would measure if we sampled $f(t)$ every T units of time apart: $f_n = ae^{-anT}$. Taking the z-transform of this sequence, it equals $az/\left(z - e^{-aT}\right)$. Recalling that $z = e^{sT} = e^{i\omega T}$, we can also plot this transform as a function of positive ω. For small ω, the transforms agree, but as ω becomes larger they diverge markedly. Why does this occur?

Recall that the z-transform is computed from a sequence comprised of samples from a continuous signal. One very important flaw in sampled data is the possible misrepresentation of high-frequency effects as lower-frequency phenomena. It is this *aliasing* or *folding* effect that we are observing here. Consequently, the z-transform of a sampled record can differ markedly from the corresponding Laplace or Fourier transforms of the continuous record at frequencies above one half of the sampling frequency. This also suggests that care should be exercised in interpolating between sampling instants. Indeed, in those applications where the output between sampling instants is very important, such as in a hybrid mixture of digital and analog systems, we must apply the so-called "modified z-transform."

Problems

From the fundamental definition of the z-transform, find the transform of the following sequences, where $n \geq 0$. Then check your answer using MATLAB.

1. $f_n = \left(\frac{1}{2}\right)^n$

2. $f_n = e^{in\theta}$

3. $f_n = \begin{cases} 1, & 0 \le n \le 5 \\ 0, & 5 < n \end{cases}$

4. $f_n = \begin{cases} \left(\frac{1}{2}\right)^n, & n = 0, 1, \ldots, 10 \\ \left(\frac{1}{4}\right)^n, & n \ge 11 \end{cases}$

5. $f_n = \begin{cases} 0, & n = 0 \\ -1, & n = 1 \\ a^n, & n \ge 2 \end{cases}$

13.2 SOME USEFUL PROPERTIES

In principle we could construct any desired transform from the definition of the z-transform. However, there are several general theorems that are much more effective in finding new transforms.

Linearity

From the definition of the z-transform, it immediately follows that

$$\text{if} \quad h_n = c_1 f_n + c_2 g_n, \quad \text{then} \quad H(z) = c_1 F(z) + c_2 G(z), \tag{13.2.1}$$

where $F(z) = \mathcal{Z}(f_n)$, $G(z) = \mathcal{Z}(g_n)$, $H(z) = \mathcal{Z}(h_n)$, and c_1, c_2 are arbitrary constants.

Multiplication by an exponential sequence

$$\text{If} \quad g_n = e^{-anT} f_n, \quad n \ge 0, \quad \text{then} \quad G(z) = F(ze^{aT}). \tag{13.2.2}$$

This follows from

$$G(z) = \mathcal{Z}(g_n) = \sum_{n=0}^{\infty} g_n\, z^{-n} = \sum_{n=0}^{\infty} e^{-anT} f_n\, z^{-n} = \sum_{n=0}^{\infty} f_n (ze^{aT})^{-n} = F(ze^{aT}). \tag{13.2.3}$$

This is the z-transform analog to the first shifting theorem in Laplace transforms.

Shifting

The effect of shifting depends upon whether it is to the right or to the left, as Table 13.2.1 illustrates. For the sequence f_{n-2}, no values from the sequence f_n are lost; thus, we anticipate that the z-transform of f_{n-2} only involves $F(z)$. However, in forming the sequence f_{n+2}, the first two values of f_n are lost, and we anticipate that the z-transform of f_{n+2} cannot be expressed solely in terms of $F(z)$ but must include those two lost pieces of information.

Table 13.2.1: Examples of Shifting Involving Sequences

n	f_n	f_{n-2}	f_{n+2}
0	1	0	4
1	2	0	8
2	4	1	16
3	8	2	64
4	16	4	128
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Let us now confirm these conjectures by finding the z-transform of f_{n+1}, which is a sequence that has been shifted one step to the left. From the definition of the z-transform, it follows that

$$\mathcal{Z}(f_{n+1}) = \sum_{n=0}^{\infty} f_{n+1} z^{-n} = z \sum_{n=0}^{\infty} f_{n+1} z^{-(n+1)} \tag{13.2.4}$$

or

$$\mathcal{Z}(f_{n+1}) = z \sum_{k=1}^{\infty} f_k z^{-k} + z f_0 - z f_0, \tag{13.2.5}$$

where we added zero in Equation 13.2.5. This algebraic trick allows us to collapse the first two terms on the right side of Equation 13.2.5 into one and

$$\mathcal{Z}(f_{n+1}) = zF(z) - zf_0. \tag{13.2.6}$$

In a similar manner, repeated applications of Equation 13.2.6 yield

$$\mathcal{Z}(f_{n+m}) = z^m F(z) - z^m f_0 - z^{m-1} f_1 - \ldots - z f_{m-1}, \tag{13.2.7}$$

where $m > 0$. This shifting operation transforms f_{n+m} into an algebraic expression involving m. Furthermore, we introduced initial sequence values, just as we introduced initial conditions when we took the Laplace transform of the nth derivative of $f(t)$. We will make frequent use of this property in solving difference equations in Section 13.4.

Consider now shifting to the right by the positive integer k,

$$g_n = f_{n-k} H_{n-k}, \quad n \geq 0, \tag{13.2.8}$$

where $H_{n-k} = 0$ for $n < k$ and 1 for $n \geq k$. Then the z-transform of Equation 13.2.8 is

$$G(z) = z^{-k} F(z), \tag{13.2.9}$$

where $G(z) = \mathcal{Z}(g_n)$, and $F(z) = \mathcal{Z}(f_n)$. This follows from

$$G(z) = \sum_{n=0}^{\infty} g_n z^{-n} = \sum_{n=0}^{\infty} f_{n-k} H_{n-k} z^{-n} = z^{-k} \sum_{n=k}^{\infty} f_{n-k} z^{-(n-k)} = z^{-k} \sum_{m=0}^{\infty} f_m z^{-m} = z^{-k} F(z). \tag{13.2.10}$$

This result is the z-transform analog to the second shifting theorem in Laplace transforms.

In symbolic calculations involving MATLAB, the operator H_{n-k} can be expressed by `Heaviside(n-k)`.

Initial-value theorem

The initial value of the sequence f_n, f_0, can be computed from $F(z)$ using the initial-value theorem:

$$f_0 = \lim_{z\to\infty} F(z). \tag{13.2.11}$$

From the definition of the z-transform,

$$F(z) = \sum_{n=0}^{\infty} f_n z^{-n} = f_0 + f_1 z^{-1} + f_2 z^{-2} + \ldots. \tag{13.2.12}$$

In the limit of $z \to \infty$, we obtain the desired result.

Final-value theorem

The value of f_n, as $n \to \infty$, is given by the final-value theorem:

$$f_\infty = \lim_{z\to 1} (z-1)F(z), \tag{13.2.13}$$

where $F(z)$ is the z-transform of f_n.

We begin by noting that

$$\mathcal{Z}(f_{n+1} - f_n) = \lim_{n\to\infty} \sum_{k=0}^{n} (f_{k+1} - f_k) z^{-k}. \tag{13.2.14}$$

Using the shifting theorem on the left side of Equation 13.2.14,

$$zF(z) - zf_0 - F(z) = \lim_{n\to\infty} \sum_{k=0}^{n} (f_{k+1} - f_k) z^{-k}. \tag{13.2.15}$$

Applying the limit as z approaches 1 to both sides of Equation 13.2.15:

$$\lim_{z\to 1} (z-1)F(z) - f_0 = \lim_{n\to\infty} \sum_{k=0}^{n} (f_{k+1} - f_k) \tag{13.2.16}$$

$$= \lim_{n\to\infty} \left[(f_1 - f_0) + (f_2 - f_1) + \ldots + (f_n - f_{n-1}) + (f_{n+1} - f_n) + \ldots\right] \tag{13.2.17}$$

$$= \lim_{n\to\infty} (-f_0 + f_{n+1}) \tag{13.2.18}$$

$$= -f_0 + f_\infty. \tag{13.2.19}$$

Consequently,

$$f_\infty = \lim_{z\to 1} (z-1)F(z). \tag{13.2.20}$$

Note that this limit has meaning only if f_∞ exists. This occurs if $F(z)$ has no second-order or higher poles on the unit circle and no poles outside the unit circle.

Multiplication by n

Given

$$g_n = nf_n, \qquad n \geq 0, \tag{13.2.21}$$

this theorem states that

$$G(z) = -z\frac{dF(z)}{dz}, \tag{13.2.22}$$

where $G(z) = \mathcal{Z}(g_n)$, and $F(z) = \mathcal{Z}(f_n)$.

This follows from

$$G(z) = \sum_{n=0}^{\infty} g_n z^{-n} = \sum_{n=0}^{\infty} nf_n z^{-n} = z\sum_{n=0}^{\infty} nf_n z^{-n-1} = -z\frac{dF(z)}{dz}. \tag{13.2.23}$$

Periodic sequence theorem

Consider the N-periodic sequence:

$$f_n = \{\underbrace{f_0 f_1 f_2 \ldots f_{N-1}}_{\text{first period}} f_0 f_1 \ldots\}, \tag{13.2.24}$$

and the related sequence:

$$x_n = \begin{cases} f_n, & 0 \leq n \leq N-1, \\ 0, & N \leq n. \end{cases} \tag{13.2.25}$$

This theorem allows us to find the z-transform of f_n if we can find the z-transform of x_n via the relationship

$$F(z) = \frac{X(z)}{1 - z^{-N}}, \qquad |z^N| > 1, \tag{13.2.26}$$

where $X(z) = \mathcal{Z}(x_n)$.

This follows from

$$F(z) = \sum_{n=0}^{\infty} f_n z^{-n} = \sum_{n=0}^{N-1} x_n z^{-n} + \sum_{n=N}^{2N-1} x_{n-N} z^{-n} + \sum_{n=2N}^{3N-1} x_{n-2N} z^{-n} + \cdots. \tag{13.2.27}$$

Application of the shifting theorem in Equation 13.2.27 leads to

$$F(z) = X(z) + z^{-N}X(z) + z^{-2N}X(z) + \cdots \tag{13.2.28}$$
$$= X(z)\left[1 + z^{-N} + z^{-2N} + \cdots\right]. \tag{13.2.29}$$

Equation 13.2.29 contains an infinite geometric series with common ratio z^{-N}, which converges if $|z^{-N}| < 1$. Thus,

$$F(z) = \frac{X(z)}{1 - z^{-N}}, \qquad |z^N| > 1. \tag{13.2.30}$$

Convolution

Given the sequences f_n and g_n, the convolution product of these two sequences is

$$w_n = f_n * g_n = \sum_{k=0}^{n} f_k g_{n-k} = \sum_{k=0}^{n} f_{n-k} g_k. \tag{13.2.31}$$

Given $F(z)$ and $G(z)$, we then have that $W(z) = F(z)G(z)$.

This follows from

$$W(z) = \sum_{n=0}^{\infty} \left[\sum_{k=0}^{n} f_k g_{n-k} \right] z^{-n} = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} f_k g_{n-k} z^{-n}, \tag{13.2.32}$$

because $g_{n-k} = 0$ for $k > n$. Reversing the order of summation and letting $m = n - k$,

$$W(z) = \sum_{k=0}^{\infty} \sum_{m=-k}^{\infty} f_k g_m z^{-(m+k)} = \left[\sum_{k=0}^{\infty} f_k z^{-k} \right] \left[\sum_{m=0}^{\infty} g_m z^{-m} \right] = F(z)G(z). \tag{13.2.33}$$

We can use MATLAB's command `conv( )`, which multiplies two polynomials to perform discrete convolution as follows:

```
>>x = [1 1 1 1 1 1 1];
>>y = [1 2 4 8 16 32 64];
>>z = conv(x,y)
```

produces

```
z =
   1 3 7 15 31 63 127 126 124 120 112 96 64
```

The first seven values of `z` contain the convolution of the sequence `x` with the sequence `y`.

Consider now the following examples of the properties discussed in this section.

• Example 13.2.1

From

$$\mathcal{Z}(a^n) = \frac{1}{1 - az^{-1}}, \tag{13.2.34}$$

for $n \geq 0$ and $|z| > |a|$, we have that

$$\mathcal{Z}\left(e^{inx}\right) = \frac{1}{1 - e^{ix} z^{-1}}, \tag{13.2.35}$$

and

$$\mathcal{Z}\left(e^{-inx}\right) = \frac{1}{1 - e^{-ix}z^{-1}}, \tag{13.2.36}$$

if $n \ge 0$ and $|z| > 1$. Therefore, the sequence $f_n = \cos(nx)$ has the z-transform

$$F(z) = \mathcal{Z}[\cos(nx)] = \tfrac{1}{2}\mathcal{Z}\left(e^{inx}\right) + \tfrac{1}{2}\mathcal{Z}\left(e^{-inx}\right) \tag{13.2.37}$$

$$= \frac{1}{2}\frac{1}{1 - e^{ix}z^{-1}} + \frac{1}{2}\frac{1}{1 - e^{-ix}z^{-1}} = \frac{1 - \cos(x)z^{-1}}{1 - 2\cos(x)z^{-1} + z^{-2}}. \tag{13.2.38}$$

□

- **Example 13.2.2**

Using the z-transform,

$$\mathcal{Z}\left(a^n\right) = \frac{1}{1 - az^{-1}}, \qquad n \ge 0, \tag{13.2.39}$$

we find that

$$\mathcal{Z}\left(na^n\right) = -z\frac{d}{dz}\left[\left(1 - az^{-1}\right)^{-1}\right] = (-z)(-1)\left(1 - az^{-1}\right)^{-2}(-a)(-1)z^{-2} \tag{13.2.40}$$

$$= \frac{az^{-1}}{\left(1 - az^{-1}\right)^2} = \frac{az}{(z-a)^2}. \tag{13.2.41}$$

□

- **Example 13.2.3**

Consider $F(z) = 2az^{-1}/(1 - az^{-1})^3$, where $|a| < |z|$ and $|a| < 1$. Here we have that

$$f_0 = \lim_{z\to\infty} F(z) = \lim_{z\to\infty} \frac{2az^{-1}}{(1 - az^{-1})^3} = 0 \tag{13.2.42}$$

from the initial-value theorem. This agrees with the inverse of $F(z)$:

$$f_n = n(n+1)a^n, \qquad n \ge 0. \tag{13.2.43}$$

If the z-transform consists of the ratio of two polynomials, we can use MATLAB to find f_0. For example, if $F(z) = 2z^2/(z-1)^3$, we can find f_0 as follows:

```
>>num = [2 0 0];
>>den = conv([1 -1],[1 -1]);
>>den = conv(den,[1 -1]);
>>initialvalue = polyval(num,1e20) / polyval(den,1e20)
initialvalue =
    2.0000e-20
```

Therefore, $f_0 = 0$. □

• Example 13.2.4

Given the z-transform $F(z) = (1-a)z/[(z-1)(z-a)]$, where $|z| > 1 > a > 0$, then from the final-value theorem we have that

$$\lim_{n\to\infty} f_n = \lim_{z\to 1}(z-1)F(z) = \lim_{z\to 1}\frac{1-a}{1-az^{-1}} = 1. \tag{13.2.44}$$

This is consistent with the inverse transform $f_n = 1 - a^n$ with $n \ge 0$. □

• Example 13.2.5

Using the sequences $f_n = 1$ and $g_n = a^n$, where a is real, verify the convolution theorem.

We first compute the convolution of f_n with g_n, namely

$$w_n = f_n * g_n = \sum_{k=0}^{n} a^k = \frac{1}{1-a} - \frac{a^{n+1}}{1-a}. \tag{13.2.45}$$

Taking the z-transform of w_n,

$$W(z) = \frac{z}{(1-a)(z-1)} - \frac{az}{(1-a)(z-a)} = \frac{z^2}{(z-1)(z-a)} = F(z)G(z), \tag{13.2.46}$$

and the convolution theorem holds true for this special case.

Problems

Use the properties of z-transforms and Table 13.1.1 to find the z-transform of the following sequences. Then check your answer using MATLAB.

1. $f_n = nTe^{-anT}$

2. $f_n = \begin{cases} 0, & n = 0 \\ na^{n-1}, & n \ge 1 \end{cases}$

3. $f_n = \begin{cases} 0, & n = 0 \\ n^2a^{n-1}, & n \ge 1 \end{cases}$

4. $f_n = a^n \cos(n)$

   $[\text{Hint : Use } \cos(n) = \frac{1}{2}(e^{in} + e^{-in})]$

5. $f_n = \cos(n-2)H_{n-2}$

6. $f_n = 3 + e^{-2nT}$

7. $f_n = \sin(n\omega_0 T + \theta)$,

8. $f_n = \begin{cases} 0, & n = 0 \\ 1, & n = 1 \\ 2, & n = 2 \\ 1, & n = 3, \end{cases} \qquad f_{n+4} = f_n$

9. $f_n = (-1)^n$

   (Hint: f_n is a periodic sequence.)

10. Using the property stated in Equation 13.2.21 and Equation 13.2.22 *twice*, find the z-transform of $n^2 = n[n(1)^n]$. Then verify your result using MATLAB.

11. Verify the convolution theorem using the sequences $f_n = g_n = 1$. Then check your results using MATLAB.

12. Verify the convolution theorem using the sequences $f_n = 1$ and $g_n = n$. Then check your results using MATLAB.

13. Verify the convolution theorem using the sequences $f_n = g_n = 1/(n!)$. Then check your results using MATLAB. Hint: Use the binomial theorem with $x = 1$ to evaluate the summation.

14. If a is a real number, show that $\mathcal{Z}(a^n f_n) = F(z/a)$, where $\mathcal{Z}(f_n) = F(z)$.

13.3 INVERSE Z-TRANSFORMS

In the previous two sections we dealt with finding the z-transform. In this section we find f_n by inverting the z-transform $F(z)$. There are four methods for finding the inverse: (1) power series, (2) recursion, (3) partial fractions, and (4) the residue method. We will discuss each technique individually. The first three apply only to those functions $F(z)$ that are *rational* functions while the residue method is more general. For symbolic computations with MATLAB, you can use `iztrans`.

Power series

By means of the long-division process, we can always rewrite $F(z)$ as the Laurent expansion:

$$F(z) = a_0 + a_1 z^{-1} + a_2 z^{-2} + \cdots. \qquad \textbf{(13.3.1)}$$

From the definition of the z-transform,

$$F(z) = \sum_{n=0}^{\infty} f_n z^{-n} = f_0 + f_1 z^{-1} + f_2 z^{-2} + \cdots, \qquad \textbf{(13.3.2)}$$

the desired sequence f_n is given by a_n.

• Example 13.3.1

Let

$$F(z) = \frac{z+1}{2z-2} = \frac{N(z)}{D(z)}. \qquad \textbf{(13.3.3)}$$

Using long division, $N(z)$ is divided by $D(z)$ and we obtain

$$F(z) = \tfrac{1}{2} + z^{-1} + z^{-2} + z^{-3} + z^{-4} + \cdots. \qquad \textbf{(13.3.4)}$$

Therefore,

$$a_0 = \tfrac{1}{2},\ a_1 = 1,\ a_2 = 1,\ a_3 = 1,\ a_4 = 1,\ \text{etc.}, \qquad \textbf{(13.3.5)}$$

which suggests that $f_0 = \frac{1}{2}$ and $f_n = 1$ for $n \geq 1$ is the inverse of $F(z)$. □

• Example 13.3.2

Let us find the inverse of the z-transform:

$$F(z) = \frac{2z^2 - 1.5z}{z^2 - 1.5z + 0.5}. \tag{13.3.6}$$

By the long-division process, we have that

$$\begin{array}{r|lllllllll}
 & 2 & + & 1.5z^{-1} & + & 1.25z^{-2} & + & 1.125z^{-3} & + & \cdots \\ \hline
z^2 - 1.5z + 0.5 & 2z^2 & - & 1.5z & & & & & & \\
 & 2z^2 & - & 3z & + & 1 & & & & \\ \cline{2-6}
 & & & 1.5z & - & 1 & & & & \\
 & & & 1.5z & - & 2.25 & + & 0.750z^{-1} & & \\ \cline{4-8}
 & & & & & 1.25 & - & 0.750z^{-1} & & \\
 & & & & & 1.25 & - & 1.870z^{-1} & + & \cdots \\ \cline{6-10}
 & & & & & & & 1.125z^{-1} & + & \cdots
\end{array}$$

Thus, $f_0 = 2$, $f_1 = 1.5$, $f_2 = 1.25$, $f_3 = 1.125$, and so forth, or $f_n = 1 + (\frac{1}{2})^n$. In general, this technique only produces numerical values for some of the elements of the sequence. Note also that our long division must always yield the power series Equation 13.3.1 in order for this method to be of any use.

To check our answer using MATLAB, we type the commands:

```
syms z; syms n positive
iztrans((2*z^2 - 1.5*z)/(z^2 - 1.5*z + 0.5),z,n)
```

which yields

```
ans =
1 + (1/2)^n
```

□

Recursive method

An alternative to long division was suggested[3] several years ago. It obtains the inverse recursively.

We begin by assuming that the z-transform is of the form

$$F(z) = \frac{a_0 z^m + a_1 z^{m-1} + a_2 z^{m-2} + \cdots + a_{m-1} z + a_m}{b_0 z^m + b_1 z^{m-1} + b_2 z^{m-2} + \cdots + b_{m-1} z + b_m}, \tag{13.3.7}$$

where some of the coefficients a_i and b_i may be zero and $b_0 \neq 0$. Applying the initial-value theorem,

$$f_0 = \lim_{z\to\infty} F(z) = a_0/b_0. \tag{13.3.8}$$

---

[3] Jury, E. I., 1964: *Theory and Application of the z-Transform Method*. John Wiley & Sons, p. 41; Pierre, D. A., 1963: A tabular algorithm for z-transform inversion. *Control Engng.*, **10(9)**, 110–111; Jenkins, L. B., 1967: A useful recursive form for obtaining inverse z-transforms. *Proc. IEEE*, **55**, 574–575.

Next, we apply the initial-value theorem to $z[F(z) - f_0]$ and find that

$$f_1 = \lim_{z\to\infty} z[F(z) - f_0] \tag{13.3.9}$$

$$= \lim_{z\to\infty} z \frac{(a_0 - b_0 f_0)z^m + (a_1 - b_1 f_0)z^{m-1} + \cdots + (a_m - b_m f_0)}{b_0 z^m + b_1 z^{m-1} + b_2 z^{m-2} + \cdots + b_{m-1}z + b_m} \tag{13.3.10}$$

$$= (a_1 - b_1 f_0)/b_0. \tag{13.3.11}$$

Note that the coefficient $a_0 - b_0 f_0 = 0$ from Equation 13.3.8. Similarly,

$$f_2 = \lim_{z\to\infty} z[zF(z) - zf_0 - f_1] \tag{13.3.12}$$

$$= \lim_{z\to\infty} z \frac{(a_0 - b_0 f_0)z^{m+1} + (a_1 - b_1 f_0 - b_0 f_1)z^m + (a_2 - b_2 f_0 - b_1 f_1)z^{m-1} + \cdots - b_m f_1}{b_0 z^m + b_1 z^{m-1} + b_2 z^{m-2} + \cdots + b_{m-1}z + b_m} \tag{13.3.13}$$

$$= (a_2 - b_2 f_0 - b_1 f_1)/b_0 \tag{13.3.14}$$

because $a_0 - b_0 f_0 = a_1 - b_1 f_0 - f_1 b_0 = 0$. Continuing this process, we finally have that

$$f_n = (a_n - b_n f_0 - b_{n-1} f_1 - \cdots - b_1 f_{n-1})/b_0, \tag{13.3.15}$$

where $a_n = b_n \equiv 0$ for $n > m$.

• Example 13.3.3

Let us redo Example 13.3.2 using the recursive method. Comparing Equation 13.3.7 to Equation 13.3.6, $a_0 = 2$, $a_1 = -1.5$, $a_2 = 0$, $b_0 = 1$, $b_1 = -1.5$, $b_2 = 0.5$, and $a_n = b_n = 0$ if $n \ge 3$. From Equation 13.3.15,

$$f_0 = a_0/b_0 = 2/1 = 2, \tag{13.3.16}$$

$$f_1 = (a_1 - b_1 f_0)/b_0 = [-1.5 - (-1.5)(2)]/1 = 1.5, \tag{13.3.17}$$

$$f_2 = (a_2 - b_2 f_0 - b_1 f_1)/b_0 \tag{13.3.18}$$

$$= [0 - (0.5)(2) - (-1.5)(1.5)]/1 = 1.25, \tag{13.3.19}$$

and

$$f_3 = (a_3 - b_3 f_0 - b_2 f_1 - b_1 f_2)/b_0 \tag{13.3.20}$$

$$= [0 - (0)(2) - (0.5)(1.5) - (-1.5)(1.25)]/1 = 1.125. \tag{13.3.21}$$

□

Partial fraction expansion

One of the popular methods for inverting Laplace transforms is partial fractions. A similar, but slightly different, scheme works here.

• **Example 13.3.4**

Given $F(z) = z/\left(z^2 - 1\right)$, let us find f_n. The first step is to obtain the partial fraction expansion of $F(z)/z$. Why we want $F(z)/z$ rather than $F(z)$ will be made clear in a moment. Thus,

$$\frac{F(z)}{z} = \frac{1}{(z-1)(z+1)} = \frac{A}{z-1} + \frac{B}{z+1}, \tag{13.3.22}$$

where

$$A = (z-1)\left.\frac{F(z)}{z}\right|_{z=1} = \frac{1}{2}, \tag{13.3.23}$$

and

$$B = (z+1)\left.\frac{F(z)}{z}\right|_{z=-1} = -\frac{1}{2}. \tag{13.3.24}$$

Multiplying Equation 13.3.22 by z,

$$F(z) = \frac{1}{2}\left(\frac{z}{z-1} - \frac{z}{z+1}\right). \tag{13.3.25}$$

Next, we find the inverse z-transform of $z/(z-1)$ and $z/(z+1)$ in Table 13.1.1. This yields

$$\mathcal{Z}^{-1}\left(\frac{z}{z-1}\right) = 1, \quad \text{and} \quad \mathcal{Z}^{-1}\left(\frac{z}{z+1}\right) = (-1)^n. \tag{13.3.26}$$

Thus, the inverse is

$$f_n = \tfrac{1}{2}\left[1 - (-1)^n\right], \; n \geq 0. \tag{13.3.27}$$

□

From this example it is clear that there are two steps: (1) obtain the partial fraction expansion of $F(z)/z$, and (2) find the inverse z-transform by referring to Table 13.1.1.

• **Example 13.3.5**

Given $F(z) = 2z^2/[(z+2)(z+1)^2]$, let us find f_n. We begin by expanding $F(z)/z$ as

$$\frac{F(z)}{z} = \frac{2z}{(z+2)(z+1)^2} = \frac{A}{z+2} + \frac{B}{z+1} + \frac{C}{(z+1)^2}, \tag{13.3.28}$$

where

$$A = (z+2)\left.\frac{F(z)}{z}\right|_{z=-2} = -4, \tag{13.3.29}$$

$$B = \left.\frac{d}{dz}\left[(z+1)^2\frac{F(z)}{z}\right]\right|_{z=-1} = 4, \tag{13.3.30}$$

and

$$C = (z+1)^2\left.\frac{F(z)}{z}\right|_{z=-1} = -2, \tag{13.3.31}$$

so that

$$F(z) = \frac{4z}{z+1} - \frac{4z}{z+2} - \frac{2z}{(z+1)^2}, \tag{13.3.32}$$

or

$$f_n = \mathcal{Z}^{-1}\left[\frac{4z}{z+1}\right] - \mathcal{Z}^{-1}\left[\frac{4z}{z+2}\right] - \mathcal{Z}^{-1}\left[\frac{2z}{(z+1)^2}\right]. \quad \textbf{(13.3.33)}$$

From Table 13.1.1,

$$\mathcal{Z}^{-1}\left(\frac{z}{z+1}\right) = (-1)^n, \quad \textbf{(13.3.34)}$$

$$\mathcal{Z}^{-1}\left(\frac{z}{z+2}\right) = (-2)^n, \quad \textbf{(13.3.35)}$$

and

$$\mathcal{Z}^{-1}\left[\frac{z}{(z+1)^2}\right] = -\ \mathcal{Z}^{-1}\left[\frac{-z}{(z+1)^2}\right] = -n(-1)^n = n(-1)^{n+1}. \quad \textbf{(13.3.36)}$$

Applying Equation 13.3.34 through Equation 13.3.36 to Equation 13.3.33,

$$f_n = 4(-1)^n - 4(-2)^n + 2n(-1)^n,\ n \geq 0. \quad \textbf{(13.3.37)}$$

□

• Example 13.3.6

Given $F(z) = (z^2+z)/(z-2)^2$, let us determine f_n. Because

$$\frac{F(z)}{z} = \frac{z+1}{(z-2)^2} = \frac{1}{z-2} + \frac{3}{(z-2)^2}, \quad \textbf{(13.3.38)}$$

$$f_n = \mathcal{Z}^{-1}\left[\frac{z}{z-2}\right] + \mathcal{Z}^{-1}\left[\frac{3z}{(z-2)^2}\right]. \quad \textbf{(13.3.39)}$$

Referring to Table 13.1.1,

$$\mathcal{Z}^{-1}\left(\frac{z}{z-2}\right) = 2^n, \quad \text{and} \quad \mathcal{Z}^{-1}\left[\frac{3z}{(z-2)^2}\right] = \tfrac{3}{2}n2^n. \quad \textbf{(13.3.40)}$$

Substituting Equation 13.3.40 into Equation 13.3.39 yields

$$f_n = \left(\tfrac{3}{2}n+1\right)2^n,\ n \geq 0. \quad \textbf{(13.3.41)}$$

□

Residue method

The power series, recursive, and partial fraction expansion methods are rather limited. We now prove that f_n may be computed from the following *inverse integral formula*:

$$f_n = \frac{1}{2\pi i}\oint_C z^{n-1}F(z)\,dz, \quad n \geq 0, \quad \textbf{(13.3.42)}$$

where C is any simple curve, taken in the positive sense, that encloses all of the singularities of $F(z)$. It is readily shown that the power series and partial fraction methods are *special cases* of the residue method.

Proof: Starting with the definition of the z-transform

$$F(z) = \sum_{n=0}^{\infty} f_n z^{-n}, \qquad |z| > R_1, \tag{13.3.43}$$

we multiply Equation 13.3.43 by z^{n-1} and integrating both sides around any contour C that includes all of the singularities:

$$\frac{1}{2\pi i} \oint_C z^{n-1} F(z)\, dz = \sum_{m=0}^{\infty} f_m \frac{1}{2\pi i} \oint_C z^{n-m} \frac{dz}{z}. \tag{13.3.44}$$

Let C be a circle of radius R, where $R > R_1$. Then, changing variables to $z = R\,e^{i\theta}$, and $dz = iz\,d\theta$,

$$\frac{1}{2\pi i} \oint_C z^{n-m} \frac{dz}{z} = \frac{R^{n-m}}{2\pi} \int_0^{2\pi} e^{i(n-m)\theta}\, d\theta = \begin{cases} 1, & m = n, \\ 0, & \text{otherwise}. \end{cases} \tag{13.3.45}$$

Substituting Equation 13.3.45 into Equation 13.3.44 yields the desired result that

$$\frac{1}{2\pi i} \oint_C z^{n-1} F(z)\, dz = f_n. \tag{13.3.46}$$

□

We can easily evaluate the inversion integral, Equation 13.3.42, using Cauchy's residue theorem.

• Example 13.3.7

Let us find the inverse z-transform of

$$F(z) = \frac{1}{(z-1)(z-2)}. \tag{13.3.47}$$

From the inversion integral,

$$f_n = \frac{1}{2\pi i} \oint_C \frac{z^{n-1}}{(z-1)(z-2)}\, dz. \tag{13.3.48}$$

Clearly the integral has simple poles at $z = 1$ and $z = 2$. However, when $n = 0$ we also have a simple pole at $z = 0$. Thus the cases $n = 0$ and $n > 0$ must be considered separately.

Case 1: $n = 0$. The residue theorem yields

$$f_0 = \text{Res}\left[\frac{1}{z(z-1)(z-2)}; 0\right] + \text{Res}\left[\frac{1}{z(z-1)(z-2)}; 1\right] + \text{Res}\left[\frac{1}{z(z-1)(z-2)}; 2\right]. \tag{13.3.49}$$

Evaluating these residues,

$$\text{Res}\left[\frac{1}{z(z-1)(z-2)}; 0\right] = \left.\frac{1}{(z-1)(z-2)}\right|_{z=0} = \frac{1}{2}, \tag{13.3.50}$$

$$\text{Res}\left[\frac{1}{z(z-1)(z-2)};1\right] = \left.\frac{1}{z(z-2)}\right|_{z=1} = -1, \tag{13.3.51}$$

and

$$\text{Res}\left[\frac{1}{z(z-1)(z-2)};2\right] = \left.\frac{1}{z(z-1)}\right|_{z=2} = \frac{1}{2}. \tag{13.3.52}$$

Substituting Equation 13.3.50 through Equation 13.3.52 into Equation 13.3.49 yields $f_0 = 0$.

Case 2: $n > 0$. Here we only have contributions from $z = 1$ and $z = 2$.

$$f_n = \text{Res}\left[\frac{z^{n-1}}{(z-1)(z-2)};1\right] + \text{Res}\left[\frac{z^{n-1}}{(z-1)(z-2)};2\right], \quad n > 0, \tag{13.3.53}$$

where

$$\text{Res}\left[\frac{z^{n-1}}{(z-1)(z-2)};1\right] = \left.\frac{z^{n-1}}{z-2}\right|_{z=1} = -1, \tag{13.3.54}$$

and

$$\text{Res}\left[\frac{z^{n-1}}{(z-1)(z-2)};2\right] = \left.\frac{z^{n-1}}{z-1}\right|_{z=2} = 2^{n-1},\ n > 0. \tag{13.3.55}$$

Thus,

$$f_n = 2^{n-1} - 1, \quad n > 0. \tag{13.3.56}$$

Combining our results,

$$f_n = \begin{cases} 0, & n = 0, \\ \frac{1}{2}\left(2^n - 2\right), & n > 0. \end{cases} \tag{13.3.57}$$

□

• Example 13.3.8

Let us use the inversion integral to find the inverse of

$$F(z) = \frac{z^2 + 2z}{(z-1)^2}. \tag{13.3.58}$$

The inversion theorem gives

$$f_n = \frac{1}{2\pi i}\oint_C \frac{z^{n+1} + 2z^n}{(z-1)^2}\,dz = \text{Res}\left[\frac{z^{n+1} + 2z^n}{(z-1)^2};1\right], \tag{13.3.59}$$

where the pole at $z = 1$ is second order. Consequently, the corresponding residue is

$$\text{Res}\left[\frac{z^{n+1} + 2z^n}{(z-1)^2};1\right] = \left.\frac{d}{dz}\left(z^{n+1} + 2z^n\right)\right|_{z=1} = 3n + 1. \tag{13.3.60}$$

Thus, the inverse z-transform of Equation 13.3.58 is

$$f_n = 3n + 1, \quad n \geq 0. \tag{13.3.61}$$

□

• Example 13.3.9

Let $F(z)$ be a z-transform whose poles lie within the unit circle $|z| = 1$. Then

$$F(z) = \sum_{n=0}^{\infty} f_n z^{-n}, \qquad |z| > 1, \tag{13.3.62}$$

and

$$F(z)F(z^{-1}) = \sum_{n=0}^{\infty} f_n^2 + \sum_{\substack{n=0 \\ n\neq m}}^{\infty} \sum_{m=0}^{\infty} f_m f_n z^{m-n}. \tag{13.3.63}$$

We now multiply both sides of Equation 13.3.63 by z^{-1} and integrate around the unit circle C. Therefore,

$$\oint_{|z|=1} F(z)F(z^{-1})z^{-1}\, dz = \sum_{n=0}^{\infty} \oint_{|z|=1} f_n^2 z^{-1}\, dz + \sum_{\substack{n=0 \\ n\neq m}}^{\infty} \sum_{m=0}^{\infty} f_m f_n \oint_{|z|=1} z^{m-n-1}\, dz, \tag{13.3.64}$$

after interchanging the order of integration and summation. Performing the integration,

$$\sum_{n=0}^{\infty} f_n^2 = \frac{1}{2\pi i} \oint_{|z|=1} F(z)F(z^{-1})z^{-1}\, dz, \tag{13.3.65}$$

which is *Parseval's theorem* for one-sided z-transforms. Recall that there are similar theorems for Fourier series and transforms. □

• Example 13.3.10: Evaluation of partial summations[4]

Consider the partial summation $S_N = \sum_{n=1}^{N} f_n$. We shall now show that z-transforms can be employed to compute S_N.

We begin by noting that

$$S_N = \sum_{n=1}^{N} f_n = \frac{1}{2\pi i} \oint_C F(z) \sum_{n=1}^{N} z^{n-1}\, dz. \tag{13.3.66}$$

Here we employed the inversion integral to replace f_n and reversed the order of integration and summation. This interchange is permissible since we only have a partial summation. Because the summation in Equation 13.3.66 is a geometric series, we have the final result that

$$S_N = \frac{1}{2\pi i} \oint_C \frac{F(z)(z^N - 1)}{z-1}\, dz. \tag{13.3.67}$$

Therefore, we can use the residue theorem and z-transforms to evaluate partial summations.

---

[4] See Bunch, K. J., W. N. Cain, and R. W. Grow, 1990: The z-transform method of evaluating partial summations in closed form. *J. Phys. A*, **23**, L1213–L1215.

Let us find $S_N = \sum_{n=1}^{N} n^3$. Because $f_n = n^3$, $F(z) = z(z^2+4z+1)/(z-1)^4$. Consequently

$$S_N = \text{Res}\left[\frac{z(z^2+4z+1)(z^N-1)}{(z-1)^5};1\right] = \frac{1}{4!}\frac{d^4}{dz^4}\left[z(z^2+4z+1)(z^N-1)\right]\Big|_{z=1} \tag{13.3.68}$$

$$= \frac{1}{4!}\frac{d^4}{dz^4}\left(z^{N+3}+4z^{N+2}+z^{N+1}-z^3-4z^2-z\right)\Big|_{z=1} = \tfrac{1}{4}(N+1)^2N^2. \tag{13.3.69}$$

□

• **Example 13.3.11**

An additional benefit of understanding inversion by the residue method is the ability to *qualitatively* anticipate the inverse by knowing the location of the poles of $F(z)$. This intuition is important because many engineering analyses discuss stability and performance entirely in terms of the properties of the system's z-transform. In Figure 13.3.1 we graphed the location of the poles of $F(z)$ and the corresponding f_n. The student should go through the mental exercise of connecting the two pictures.

Problems

Use the power series or recursive method to compute the first few values of f_n of the following z-transforms. Then check your answers with MATLAB.

1. $F(z) = \dfrac{0.09z^2+0.9z+0.09}{12.6z^2-24z+11.4}$

2. $F(z) = \dfrac{z+1}{2z^4-2z^3+2z-2}$

3. $F(z) = \dfrac{1.5z^2+1.5z}{15.25z^2-36.75z+30.75}$

4. $F(z) = \dfrac{6z^2+6z}{19z^3-33z^2+21z-7}$

Use partial fractions to find the inverse of the following z-transforms. Then verify your answers with MATLAB.

5. $F(z) = \dfrac{z(z+1)}{(z-1)(z^2-z+1/4)}$

6. $F(z) = \dfrac{(1-e^{-aT})z}{(z-1)(z-e^{-aT})}$

7. $F(z) = \dfrac{z^2}{(z-1)(z-\alpha)}$

8. $F(z) = \dfrac{(2z-a-b)z}{(z-a)(z-b)}$

9. Using the property that the z-transform of $g_n = f_{n-k}H_{n-k}$ if $n \ge 0$ is $G(z) = z^{-k}F(z)$, find the inverse of

$$F(z) = \frac{z+1}{z^{10}(z-1/2)}.$$

Then check your answer with MATLAB.

Use the residue method to find the inverse z-transform of the following z-transforms. Then verify your answer with MATLAB.

10. $F(z) = \dfrac{z^2+3z}{(z-1/2)^3}$

11. $F(z) = \dfrac{z}{(z+1)^2(z-2)}$

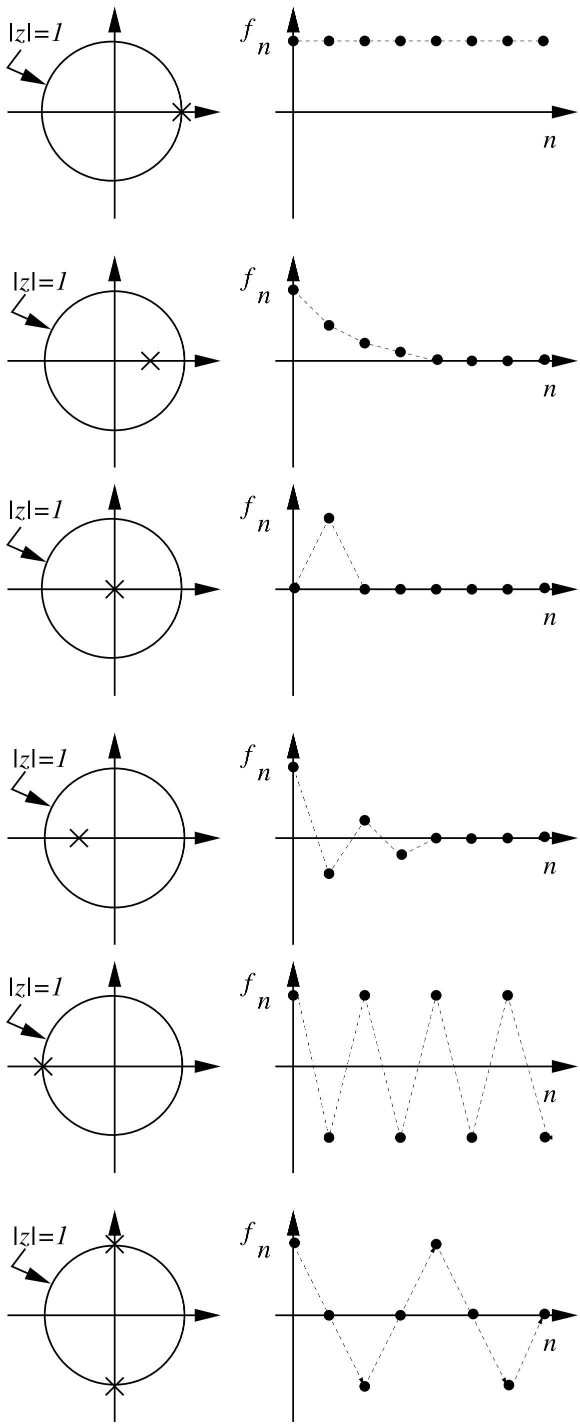


Figure 13.3.1: The correspondence between the location of the simple poles of the z-transform $F(z)$ and the behavior of f_n.

12. $F(z) = \dfrac{z}{(z+1)^2(z-1)^2}$

13. $F(z) = e^{a/z}$

13.4 SOLUTION OF DIFFERENCE EQUATIONS

Having reached the point where we can take a z-transform and then find its inverse, we are ready to use it to solve difference equations. The procedure parallels that of solving ordinary differential equations by Laplace transforms. Essentially we reduce the difference equation to an algebraic problem. We then find the solution by inverting $Y(z)$.

• Example 13.4.1

Let us solve the second-order difference equation

$$2y_{n+2} - 3y_{n+1} + y_n = 5\ 3^n,\ \ n \geq 0, \tag{13.4.1}$$

where $y_0 = 0$ and $y_1 = 1$.

Taking the z-transform of both sides of Equation 13.4.1, we obtain

$$2\mathcal{Z}(y_{n+2}) - 3\mathcal{Z}(y_{n+1}) + \mathcal{Z}(y_n) = 5\,\mathcal{Z}(3^n). \tag{13.4.2}$$

From the shifting theorem and Table 13.1.1,

$$2z^2Y(z) - 2z^2y_0 - 2zy_1 - 3[zY(z) - zy_0] + Y(z) = \frac{5z}{z-3}. \tag{13.4.3}$$

Substituting $y_0 = 0$ and $y_1 = 1$ into Equation 13.4.3 and simplifying yields

$$(2z-1)(z-1)Y(z) = \frac{z(2z-1)}{z-3}, \tag{13.4.4}$$

or

$$Y(z) = \frac{z}{(z-3)(z-1)}. \tag{13.4.5}$$

To obtain y_n from $Y(z)$ we can employ partial fractions or the residue method. Applying partial fractions gives

$$\frac{Y(z)}{z} = \frac{A}{z-1} + \frac{B}{z-3}, \tag{13.4.6}$$

where

$$A = (z-1)\left.\frac{Y(z)}{z}\right|_{z=1} = -\frac{1}{2}, \tag{13.4.7}$$

and

$$B = (z-3)\left.\frac{Y(z)}{z}\right|_{z=3} = \frac{1}{2}. \tag{13.4.8}$$

Thus,

$$Y(z) = -\frac{1}{2}\frac{z}{z-1} + \frac{1}{2}\frac{z}{z-3}, \tag{13.4.9}$$

or

$$y_n = -\frac{1}{2}\mathcal{Z}^{-1}\left(\frac{z}{z-1}\right) + \frac{1}{2}\mathcal{Z}^{-1}\left(\frac{z}{z-3}\right). \tag{13.4.10}$$

From Equation 13.4.10 and Table 13.1.1,

$$y_n = \tfrac{1}{2}\left(3^n - 1\right), \qquad n \geq 0. \tag{13.4.11}$$

An alternative to this hand calculation is to use MATLAB's `ztrans` and `iztrans` to solve difference equations. In the present case, the MATLAB script would read

```
clear
% define symbolic variables
syms z Y; syms n positive
% take z-transform of left side of difference equation
LHS = ztrans(2*sym('y(n+2)')-3*sym('y(n+1)')+sym('y(n)'),n,z);
% take z-transform of right side of difference equation
RHS = 5 * ztrans(3^n,n,z);
% set Y for z-transform of y and introduce initial conditions
newLHS = subs(LHS,'ztrans(y(n),n,z)','y(0)','y(1)',Y,0,1);
% solve for Y
Y = solve(newLHS-RHS,Y);
% invert z-transform and find y(n)
y = iztrans(Y,z,n)
```

This script produced

```
y =
-1/2+1/2*3^n
```

Two checks confirm that we have the *correct* solution. First, our solution must satisfy the initial values of the sequence. Computing y_0 and y_1,

$$y_0 = \tfrac{1}{2}(3^0 - 1) = \tfrac{1}{2}(1 - 1) = 0, \tag{13.4.12}$$

and

$$y_1 = \tfrac{1}{2}(3^1 - 1) = \tfrac{1}{2}(3 - 1) = 1. \tag{13.4.13}$$

Thus, our solution gives the correct initial values.

Our sequence y_n must also satisfy the difference equation. Now

$$y_{n+2} = \tfrac{1}{2}(3^{n+2} - 1) = \tfrac{1}{2}(9\ 3^n - 1), \tag{13.4.14}$$

and

$$y_{n+1} = \tfrac{1}{2}(3^{n+1} - 1) = \tfrac{1}{2}(3\ 3^n - 1). \tag{13.4.15}$$

Therefore,

$$2y_{n+2} - 3y_{n+1} + y_n = \left(9 - \tfrac{9}{2} + \tfrac{1}{2}\right) 3^n - 1 + \tfrac{3}{2} - \tfrac{1}{2} = 5\ 3^n \tag{13.4.16}$$

and our solution is correct.

Finally, we note that the term $3^n/2$ is necessary to give the right side of Equation 13.4.1; it is the particular solution. The $-1/2$ term is necessary so that the sequence satisfies the initial values; it is the complementary solution.

□

• **Example 13.4.2**

Let us find the y_n in the difference equation

$$y_{n+2} - 2y_{n+1} + y_n = 1, \quad n \ge 0 \tag{13.4.17}$$

with the initial conditions $y_0 = 0$ and $y_1 = 3/2$.

From Equation 13.4.17,

$$\mathcal{Z}(y_{n+2}) - 2\mathcal{Z}(y_{n+1}) + \mathcal{Z}(y_n) = \mathcal{Z}(1). \tag{13.4.18}$$

The z-transform of the left side of Equation 13.4.18 is obtained from the shifting theorem, and Table 13.1.1 yields $\mathcal{Z}(1)$. Thus,

$$z^2Y(z) - z^2y_0 - zy_1 - 2zY(z) + 2zy_0 + Y(z) = \frac{z}{z-1}. \tag{13.4.19}$$

Substituting $y_0 = 0$ and $y_1 = 3/2$ in Equation 13.4.19 and simplifying gives

$$Y(z) = \frac{3z^2 - z}{2(z-1)^3} \tag{13.4.20}$$

or

$$y_n = \mathcal{Z}^{-1}\left[\frac{3z^2 - z}{2(z-1)^3}\right]. \tag{13.4.21}$$

We find the inverse z-transform of Equation 13.4.21 by the residue method, or

$$y_n = \frac{1}{2\pi i}\oint_C \frac{3z^{n+1} - z^n}{2(z-1)^3}\,dz = \frac{1}{2!}\frac{d^2}{dz^2}\left[\frac{3z^{n+1}}{2} - \frac{z^n}{2}\right]\Bigg|_{z=1} \tag{13.4.22}$$

$$= \tfrac{1}{2}n^2 + n. \tag{13.4.23}$$

Thus,

$$y_n = \tfrac{1}{2}n^2 + n, \quad n \ge 0. \tag{13.4.24}$$

Note that $n^2/2$ gives the particular solution to Equation 13.4.17, while n is there so that y_n satisfies the initial conditions. This problem is particularly interesting because our constant forcing produces a response that grows as n^2, just as in the case of resonance in a time-continuous system when a finite forcing such as $\sin(\omega_0 t)$ results in a response whose amplitude grows as t^m. □

• **Example 13.4.3**

Let us solve the difference equation

$$b^2y_n + y_{n+2} = 0, \tag{13.4.25}$$

where the initial conditions are $y_0 = b^2$ and $y_1 = 0$.

We begin by taking the z-transform of each term in Equation 13.4.25. This yields

$$b^2\mathcal{Z}(y_n) + \mathcal{Z}(y_{n+2}) = 0. \tag{13.4.26}$$

From the shifting theorem, it follows that

$$b^2Y(z) + z^2Y(z) - z^2y_0 - zy_1 = 0. \tag{13.4.27}$$

Substituting $y_0 = b^2$ and $y_1 = 0$ into Equation 13.4.27,

$$b^2Y(z) + z^2Y(z) - b^2z^2 = 0, \tag{13.4.28}$$

or

$$Y(z) = \frac{b^2z^2}{z^2+b^2}. \tag{13.4.29}$$

To find y_n we employ the residue method or

$$y_n = \frac{1}{2\pi i}\oint_C \frac{b^2z^{n+1}}{(z-ib)(z+ib)}\,dz. \tag{13.4.30}$$

Thus,

$$y_n = \left.\frac{b^2z^{n+1}}{z+ib}\right|_{z=ib} + \left.\frac{b^2z^{n+1}}{z-ib}\right|_{z=-ib} = \frac{b^{n+2}i^n}{2} + \frac{b^{n+2}(-i)^n}{2} \tag{13.4.31}$$

$$= \frac{b^{n+2}e^{in\pi/2}}{2} + \frac{b^{n+2}e^{-in\pi/2}}{2} = b^{n+2}\cos\left(\frac{n\pi}{2}\right), \tag{13.4.32}$$

because $\cos(x) = \frac{1}{2}\left(e^{ix}+e^{-ix}\right)$. Consequently, we obtain the desired result that

$$y_n = b^{n+2}\cos\left(\frac{n\pi}{2}\right) \quad \text{for } n \ge 0. \tag{13.4.33}$$

□

• Example 13.4.4: Compound interest

Difference equations arise in finance because the increase or decrease in an account occurs in discrete steps. For example, the amount of money in a compound interest savings account after $n+1$ conversion periods (the time period between interest payments) is

$$y_{n+1} = y_n + ry_n, \tag{13.4.34}$$

where r is the interest rate per conversion period. The second term on the right side of Equation 13.4.34 is the amount of interest paid at the end of each period.

Let us ask a somewhat more difficult question of how much money we will have if we withdraw the amount A at the end of every period starting after the period ℓ. Now the difference equation reads

$$y_{n+1} = y_n + ry_n - AH_{n-\ell-1}. \tag{13.4.35}$$

Taking the z-transform of Equation 13.4.35,

$$zY(z) - zy_0 = (1+r)Y(z) - \frac{Az^{2-\ell}}{z-1} \tag{13.4.36}$$

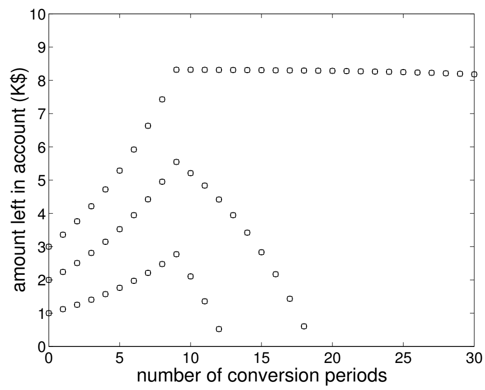


Figure 13.4.1: The amount in a savings account as a function of an annual conversion period when interest is compounded at the annual rate of 12% and \$1000 is taken from the account every period starting with period 10.

after using Equation 13.2.9 or

$$Y(z) = \frac{y_0 z}{z - (1 + r)} - \frac{A z^{2-\ell}}{(z - 1)[z - (1 + r)]}. \tag{13.4.37}$$

Taking the inverse of Equation 13.4.37,

$$y_n = y_0(1 + r)^n - \frac{A}{r}\left[(1 + r)^{n-\ell+1} - 1\right] H_{n-\ell}. \tag{13.4.38}$$

The first term in Equation 13.4.38 represents the growth of money by compound interest while the second term gives the depletion of the account by withdrawals.

Figure 13.4.1 gives the values of y_n for various starting amounts assuming an annual conversion period with $r = 0.12$, $\ell = 10$ years, and $A = \$1000$. These computations were done two ways using MATLAB as follows:

```
% load in parameters
clear; r = 0.12; A = 1; k = 0:30;
y = zeros(length(k),3); yanswer = zeros(length(k),3);
% set initial condition
for m=1:3
  y(1,m) = m;
% compute other y values
  for n = 1:30
    y(n+1,m) = y(n,m)+r*y(n,m);
    y(n+1,m) = y(n+1,m)-A*stepfun(n,11);
  end
% now use Equation 13.4.38
  for n = 1:31
    yanswer(n,m) = y(1,m)*(1+r)^(n-1);
    yanswer(n,m) = yanswer(n,m)-A*((1+r)^(n-10)-1)
            *stepfun(n,11)/r;
end; end;
plot(k,y,'o'); hold; plot(k,yanswer,'s');
```

```
axis([0 30 0 10])
xlabel('number of conversion periods','Fontsize',20)
ylabel('amount left in account (K$)','Fontsize',20)
```

Figure 13.4.1 shows that if an investor places an initial amount of \$3000 in an account bearing 12% annually, after 10 years he can withdraw \$1000 annually, essentially forever. This is because the amount that he removes every year is replaced by the interest on the funds that remain in the account.

□

• **Example 13.4.5**

Let us solve the following system of difference equations:

$$x_{n+1} = 4x_n + 2y_n, \tag{13.4.39}$$

and

$$y_{n+1} = 3x_n + 3y_n, \tag{13.4.40}$$

with the initial values of $x_0 = 0$ and $y_0 = 5$.

Taking the z-transform of Equation 13.4.39 and Equation 13.4.40,

$$zX(z) - x_0 z = 4X(z) + 2Y(z), \tag{13.4.41}$$

$$zY(z) - y_0 z = 3X(z) + 3Y(z), \tag{13.4.42}$$

or

$$(z-4)X(z) - 2Y(z) = 0, \tag{13.4.43}$$

$$3X(z) - (z-3)Y(z) = -5z. \tag{13.4.44}$$

Solving for $X(z)$ and $Y(z)$,

$$X(z) = \frac{10z}{(z-6)(z-1)} = \frac{2z}{z-6} - \frac{2z}{z-1}, \tag{13.4.45}$$

and

$$Y(z) = \frac{5z(z-4)}{(z-6)(z-1)} = \frac{2z}{z-6} + \frac{3z}{z-1}. \tag{13.4.46}$$

Taking the inverse of Equation 13.4.45 and Equation 13.4.46 term by term,

$$x_n = -2 + 2\ 6^n, \qquad \text{and} \qquad y_n = 3 + 2\ 6^n. \tag{13.4.47}$$

We can also check our work using the MATLAB script

```
clear
% define symbolic variables
syms X Y z; syms n positive
% take z-transform of left side of differential equations
LHS1 = ztrans(sym('x(n+1)')-4*sym('x(n)')-2*sym('y(n)'),n,z);
LHS2 = ztrans(sym('y(n+1)')-3*sym('x(n)')-3*sym('y(n)'),n,z);
% set X and Y for the z-transform of x and y
%     and introduce initial conditions
```

```
newLHS1 = subs(LHS1,'ztrans(x(n),n,z)','ztrans(y(n),n,z)',...
        'x(0)','y(0)',X,Y,0,5);
newLHS2 = subs(LHS2,'ztrans(x(n),n,z)','ztrans(y(n),n,z)',...
        'x(0)','y(0)',X,Y,0,5);
% solve for X and Y
[X,Y] = solve(newLHS1,newLHS2,X,Y);
% invert z-transform and find x(n) and y(n)
x = iztrans(X,z,n)
y = iztrans(Y,z,n)
```

This script yields

```
x =
2*6^n-2
y =
2*6^n+3
```

Problems

Solve the following difference equations using z-transforms, where $n \geq 0$. Check your answer using MATLAB.

1. $y_{n+1} - y_n = n^2, \quad y_0 = 1.$
2. $y_{n+2} - 2y_{n+1} + y_n = 0, \quad y_0 = y_1 = 1.$
3. $y_{n+2} - 2y_{n+1} + y_n = 1, \quad y_0 = y_1 = 0.$
4. $y_{n+1} + 3y_n = n, \quad y_0 = 0.$
5. $y_{n+1} - 5y_n = \cos(n\pi), \quad y_0 = 0.$
6. $y_{n+2} - 4y_n = 1, \quad y_0 = 1, y_1 = 0.$
7. $y_{n+2} - \frac{1}{4}y_n = (\frac{1}{2})^n, y_0 = y_1 = 0.$
8. $y_{n+2} - 5y_{n+1} + 6y_n = 0, y_0 = y_1 = 1.$
9. $y_{n+2} - 3y_{n+1} + 2y_n = 1, \quad y_0 = y_1 = 0.$
10. $y_{n+2} - 2y_{n+1} + y_n = 2, \quad y_0 = 0, y_1 = 2.$
11. $x_{n+1} = 3x_n - 4y_n,\ y_{n+1} = 2x_n - 3y_n, \quad x_0 = 3,\ y_0 = 2.$
12. $x_{n+1} = 2x_n - 10y_n,\ y_{n+1} = -x_n - y_n, \quad x_0 = 3,\ y_0 = -2.$
13. $x_{n+1} = x_n - 2y_n,\ y_{n+1} = -6y_n, \quad x_0 = -1,\ y_0 = -7.$
14. $x_{n+1} = 4x_n - 5y_n,\ y_{n+1} = x_n - 2y_n, \quad x_0 = 6,\ y_0 = 2.$

13.5 STABILITY OF DISCRETE-TIME SYSTEMS

When we discussed the solution of ordinary differential equations by Laplace transforms, we introduced the concept of transfer function and impulse response. In the case of discrete-time systems, similar considerations come into play.

Consider the recursive system

$$y_n = a_1 y_{n-1} H_{n-1} + a_2 y_{n-2} H_{n-2} + x_n, \quad n \geq 0, \tag{13.5.1}$$

where H_{n-k} is the unit step function. It equals 0 for $n < k$ and 1 for $n \geq k$. Equation 13.5.1 is called a *recursive system* because future values of the sequence depend upon all of the previous values. At present, a_1 and a_2 are free parameters that we shall vary.

Using Equation 13.2.7,

$$z^2Y(z) - a_1zY(z) - a_2Y(z) = z^2X(z), \tag{13.5.2}$$

or

$$G(z) = \frac{Y(z)}{X(z)} = \frac{z^2}{z^2 - a_1z - a_2}. \tag{13.5.3}$$

As in the case of Laplace transforms, the ratio $Y(z)/X(z)$ is the transfer function. The inverse of the transfer function gives the impulse response for our discrete-time system. This particular transfer function has two poles, namely

$$z_{1,2} = \frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} + a_2}. \tag{13.5.4}$$

At this point, we consider three cases.

Case 1: $a_1^2/4 + a_2 < 0$. In this case z_1 and z_2 are complex conjugates. Let us write them as $z_{1,2} = re^{\pm i\omega_0 T}$. Then

$$G(z) = \frac{z^2}{(z - re^{i\omega_0 T})(z - re^{-i\omega_0 T})} = \frac{z^2}{z^2 - 2r\cos(\omega_0 T)z + r^2}, \tag{13.5.5}$$

where $r^2 = -a_2$, and $\omega_0 T = \cos^{-1}(a_1/2r)$. From the inversion integral,

$$g_n = \operatorname{Res}\left[\frac{z^{n+1}}{z^2 - 2r\cos(\omega_0 T)z + r^2}; z_1\right] + \operatorname{Res}\left[\frac{z^{n+1}}{z^2 - 2r\cos(\omega_0 T)z + r^2}; z_2\right], \tag{13.5.6}$$

where g_n denotes the impulse response. Now

$$\operatorname{Res}\left[\frac{z^{n+1}}{z^2 - 2r\cos(\omega_0 T)z + r^2}; z_1\right] = \lim_{z\to z_1} \frac{(z - z_1)z^{n+1}}{(z - z_1)(z - z_2)} \tag{13.5.7}$$

$$= r^n \frac{\exp[i(n+1)\omega_0 T]}{e^{i\omega_0 T} - e^{-i\omega_0 T}} \tag{13.5.8}$$

$$= \frac{r^n \exp[i(n+1)\omega_0 T]}{2i\sin(\omega_0 T)}. \tag{13.5.9}$$

Similarly,

$$\operatorname{Res}\left[\frac{z^{n+1}}{z^2 - 2r\cos(\omega_0 T)z + r^2}; z_2\right] = -\frac{r^n \exp[-i(n+1)\omega_0 T]}{2i\sin(\omega_0 T)}, \tag{13.5.10}$$

and

$$g_n = \frac{r^n \sin[(n+1)\omega_0 T]}{\sin(\omega_0 T)}. \tag{13.5.11}$$

A graph of $\sin[(n+1)\omega_0 T]/\sin(\omega_0 T)$ with respect to n gives a sinusoidal envelope. More importantly, if $|r| < 1$ these oscillations vanish as $n \to \infty$ and the system is stable. On the other hand, if $|r| > 1$ the oscillations grow without bound as $n \to \infty$ and the system is unstable.

Recall that $|r| > 1$ corresponds to poles that lie outside the unit circle while $|r| < 1$ is exactly the opposite. Our example suggests that for discrete-time systems to be stable, all of the poles of the transfer function must lie within the unit circle while an unstable system has at least one pole that lies outside of this circle.

Case 2: $a_1^2/4 + a_2 > 0$. This case leads to two real roots, z_1 and z_2. From the inversion integral, the sum of the residues gives the impulse response

$$g_n = \frac{z_1^{n+1} - z_2^{n+1}}{z_1 - z_2}. \tag{13.5.12}$$

Once again, if the poles lie within the unit circle, $|z_1| < 1$ and $|z_2| < 1$, the system is stable.

Case 3: $a_1^2/4 + a_2 = 0$. This case yields $z_1 = z_2$,

$$G(z) = \frac{z^2}{(z - a_1/2)^2} \tag{13.5.13}$$

and

$$g_n = \frac{1}{2\pi i} \oint_C \frac{z^{n+1}}{(z - a_1/2)^2} dz = \left(\frac{a_1}{2}\right)^n (n+1). \tag{13.5.14}$$

This system is obviously stable if $|a_1/2| < 1$ and the pole of the transfer function lies within the unit circle.

In summary, finding the transfer function of a discrete-time system is important in determining its stability. Because the location of the poles of $G(z)$ determines the response of the system, a stable system has all of its poles within the unit circle. Conversely, if any of the poles of $G(z)$ lie outside of the unit circle, the system is unstable. Finally, if $\lim_{n\to\infty} g_n = c$, the system is marginally stable. For example, if $G(z)$ has simple poles, some of the poles must lie *on* the unit circle.

• Example 13.5.1

Numerical methods of integration provide some of the simplest, yet most important, difference equations in the literature. In this example,[5] we show how z-transforms can be used to highlight the strengths and weaknesses of such schemes.

Consider the trapezoidal integration rule in numerical analysis. The integral y_n is updated by adding the latest trapezoidal approximation of the continuous curve. Thus, the integral is computed by

$$y_n = \tfrac{1}{2} T(x_n + x_{n-1} H_{n-1}) + y_{n-1} H_{n-1}, \tag{13.5.15}$$

where T is the interval between evaluations of the integrand.

We first determine the stability of this rule because it is of little value if it is not stable. Using Equation 13.2.7, the transfer function is

$$G(z) = \frac{Y(z)}{X(z)} = \frac{T}{2}\left(\frac{z+1}{z-1}\right). \tag{13.5.16}$$

To find the impulse response, we use the inversion integral and find that

$$g_n = \frac{T}{4\pi i} \oint_C z^{n-1} \frac{z+1}{z-1} dz. \tag{13.5.17}$$

[5] See Salzer, J. M., 1954: Frequency analysis of digital computers operating in real time. *Proc. IRE*, **42**, 457–466.

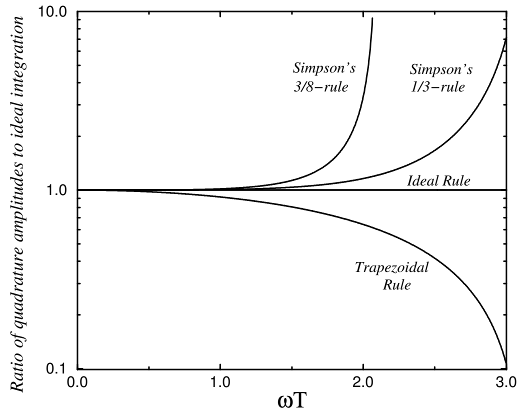


Figure 13.5.1: Comparison of various quadrature formulas by ratios of their amplitudes to that of an ideal integrator. (From Salzer, J. M., 1954: Frequency analysis of digital computers operating in real time. *Proc. IRE*, **42**, p. 463.)

At this point, we must consider two cases: $n = 0$ and $n > 0$. For $n = 0$,

$$g_0 = \frac{T}{2}\operatorname{Res}\left[\frac{z+1}{z(z-1)}; 0\right] + \frac{T}{2}\operatorname{Res}\left[\frac{z+1}{z(z-1)}; 1\right] = \frac{T}{2}. \quad \textbf{(13.5.18)}$$

For $n > 0$,

$$g_0 = \frac{T}{2}\operatorname{Res}\left[\frac{z^{n-1}(z+1)}{z-1}; 1\right] = T. \quad \textbf{(13.5.19)}$$

Therefore, the impulse response for this numerical scheme is $g_0 = \frac{T}{2}$ and $g_n = T$ for $n > 0$. Note that this is a marginally stable system (the solution neither grows nor decays with n) because the pole associated with the transfer function lies *on* the unit circle.

Having discovered that the system is not unstable, let us continue and explore some of its properties. Recall now that $z = e^{sT} = e^{i\omega T}$ if $s = i\omega$. Then the transfer function becomes

$$G(\omega) = \frac{T}{2}\frac{1+e^{-i\omega T}}{1-e^{-i\omega T}} = -\frac{iT}{2}\cot\left(\frac{\omega T}{2}\right). \quad \textbf{(13.5.20)}$$

On the other hand, the transfer function of an ideal integrator is $1/s$ or $-i/\omega$. Thus, the trapezoidal rule has ideal phase but its shortcoming lies in its amplitude characteristic; it lies below the ideal integrator for $0 < \omega T < \pi$. We show this behavior, along with that for Simpson's one-third rule and Simpson's three-eighths rule, in Figure 13.5.1.

Figure 13.5.1 confirms the superiority of Simpson's one-third rule over his three-eighths rule. The figure also shows that certain schemes are better at suppressing noise at higher frequencies, an effect not generally emphasized in numerical calculus but often important in system design. For example, the trapezoidal rule is inferior to all others at low frequencies but only to Simpson's one-third rule at higher frequencies. Furthermore, the trapezoidal rule might actually be preferred, not only because of its simplicity but also because it attenuates at higher frequencies, thereby counteracting the effect of noise. □

• Example 13.5.2

Given the transfer function

$$G(z) = \frac{z^2}{(z-1)(z-1/2)}, \quad \textbf{(13.5.21)}$$

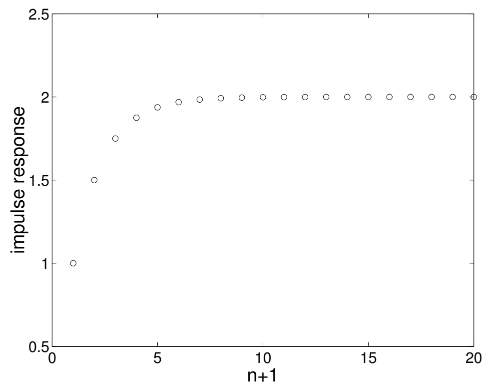


Figure 13.5.2: The impulse response for a discrete system with a transform function given by Equation 13.5.21.

is this discrete-time system stable or marginally stable?

This transfer function has two simple poles. The pole at $z = 1/2$ gives rise to a term that varies as $(\frac{1}{2})^n$ in the impulse response while the $z = 1$ pole gives a constant. Because this constant neither grows nor decays with n, the system is marginally stable. □

• Example 13.5.3

In most cases the transfer function consists of a ratio of two polynomials. In this case we can use the MATLAB function `filter` to compute the impulse response as follows: Consider the Kronecker delta sequence, $x_0 = 1$, and $x_n = 0$ for $n > 0$. From the definition of the z-transform, $X(z) = 1$. Therefore, if our input into `filter` is the Kronecker delta sequence, the output y_n will be the impulse response since $Y(z) = G(z)$. If the impulse response grows without bound as n increases, the system is unstable. If it goes to zero as n increases, the system is stable. If it remains constant, it is marginally stable.

To illustrate this concept, the following MATLAB script finds the impulse response corresponding to the transfer function, Equation 13.5.21:

```
% enter the coefficients of the numerator
%     of the transfer function, Equation 13.5.21
num = [1 0 0];
% enter the coefficients of the denominator
%     of the transfer function, Equation 13.5.21
den = [1 -1.5 0.5];
% create the Kronecker delta sequence
x = [1 zeros(1,20)];
% find the impulse response
y = filter(num,den,x);
% plot impulse response
plot(y,'o'), axis([0 20 0.5 2.5])
xlabel('n+1','Fontsize',20)
ylabel('impulse response','Fontsize',20)
```

Figure 13.5.2 shows the computed impulse response. The asymptotic limit is two, so the system is marginally stable as we found before.

We note in closing that the same procedure can be used to find the inverse of *any* z-transform that consists of a ratio of two polynomials. Here we simply set $G(z)$ equal to the given z-transform and perform the same analysis.

Problems

For the following time-discrete systems, find the transfer function and determine whether the systems are unstable, marginally stable, or stable. Check your answer by graphing the impulse response using MATLAB.

1. $y_n = y_{n-1}H_{n-1} + x_n$
2. $y_n = 2y_{n-1}H_{n-1} - y_{n-2}H_{n-2} + x_n$
3. $y_n = 3y_{n-1}H_{n-1} + x_n$
4. $y_n = \frac{1}{4}y_{n-2}H_{n-2} + x_n$

Further Readings

Jury, E. I., 1964: *Theory and Application of the z-Transform Method*. John Wiley & Sons, 330 pp. The classic text on z-transforms.

LePage, W. R., 1980: *Complex Variables and the Laplace Transform for Engineers.* Dover, 483 pp. Chapter 16 is on z-transforms.

Chapter 14
The Hilbert Transform

In addition to the Fourier, Laplace, and z-transforms, there are many other linear transforms that have their own special niche in engineering. Examples include Hankel, Walsh, Radon, and Hartley transforms. In this chapter we consider the *Hilbert transform*, which is a commonly used technique for relating the real and imaginary parts of a spectral response, particularly in communication theory.

We begin our study of Hilbert transforms by first defining them and then exploring their properties. Next, we develop the concept of the analytic signal. Finally, we explore a property of Hilbert transforms that is frequently applied to data analysis: the Kramers-Kronig relationship.

14.1 DEFINITION

In Chapter 13 we motivated the development of z-transforms by exploring the concept of the ideal sampler. In the case of Hilbert transforms, we introduce another fundamental operation, namely *quadrature phase shifting* or the *ideal Hilbert transformer*. This procedure does nothing more than shift the phase of all input frequency components by $-\pi/2$. Hilbert transformers are frequently used in communication systems and signal processing; examples include the generation of single-sideband modulated signals and radar and speech signal processing.

Because a $-\pi/2$ phase shift is equivalent to multiplying the Fourier transform of a signal by $e^{-i\pi/2} = -i$, and because phase shifting must be an odd function of frequency,[1]

[1] For a real function the phase of its Fourier transform must be an odd function of ω.

the transfer function of the phase shifter is $G(\omega) = -i\ \text{sgn}(\omega)$, where $\text{sgn}(\cdot)$ is defined by Equation 11.2.11. In other words, if $X(\omega)$ denotes the input spectrum to the phase shifter, the output spectrum must be $-i\ \text{sgn}(\omega)X(\omega)$. If the process is repeated, the total phase shift is $-\pi$, a complete phase reversal of all frequency components. The output spectrum then equals $[-i\ \text{sgn}(\omega)]^2 X(\omega) = -X(\omega)$. This agrees with the notion of phase reversal because the output function is $-x(t)$.

Consider now the impulse response of the quadrature phase shifter, $g(t) = \mathcal{F}^{-1}[G(\omega)]$. From the definition of Fourier transforms,

$$\frac{dG}{d\omega} = -i \int_{-\infty}^{\infty} t g(t) e^{-i\omega t}\, dt, \tag{14.1.1}$$

and

$$g(t) = \frac{i}{t} \mathcal{F}^{-1}\left(\frac{dG}{d\omega}\right). \tag{14.1.2}$$

Since $G'(\omega) = -2i\delta(\omega)$, the corresponding impulse response is

$$g(t) = \frac{i}{t} \mathcal{F}^{-1}[-2i\delta(\omega)] = \frac{1}{\pi t}. \tag{14.1.3}$$

Consequently, if $x(t)$ is the input to a quadrature phase shifter, the superposition integral gives the output time function as

$$\boxed{\widehat{x}(t) = x(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau}\, d\tau.} \tag{14.1.4}$$

We shall define $\widehat{x}(t)$ as the *Hilbert transform* of $x(t)$, although some authors use the negative of Equation 14.1.4 corresponding to a $+\pi/2$ phase shift. The transform $\widehat{x}(t)$ is also called the *harmonic conjugate* of $x(t)$.

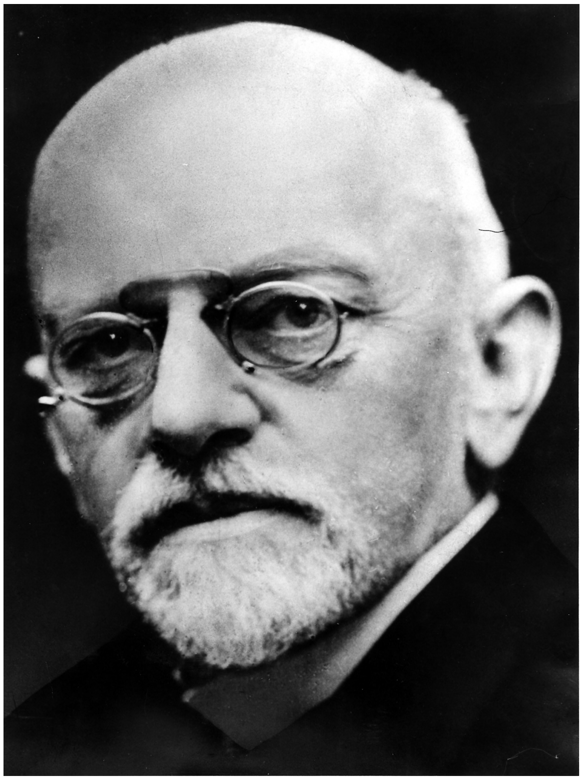
In similar fashion, $\widehat{\widehat{x}}(t)$ is the Hilbert transform of the Hilbert transform of $x(t)$ and corresponds to the output of two cascaded phase shifters. However, this output is known to be $-x(t)$, so $\widehat{\widehat{x}}(t) = -x(t)$, and we arrive at the *inverse Hilbert transform* relationship that

$$\boxed{x(t) = -\widehat{x}(t) * \frac{1}{\pi t} = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\widehat{x}(\tau)}{t - \tau}\, d\tau.} \tag{14.1.5}$$

Taken together, $x(t)$ and $\widehat{x}(t)$ are called a *Hilbert pair*. Hilbert pairs enjoy the unique property that $x(t) + i\widehat{x}(t)$ is an *analytic function*.[2]

---

[2] For the proof, see Titchmarsh, E. C., 1948: *Introduction to the Theory of Fourier Integrals*. Oxford University Press, p. 125.



Descended from a Prussian middle-class family, David Hilbert (1862–1943) would make significant contributions in the fields of algebraic form, algebraic number theory, foundations of geometry, analysis, mathematical physics, and the foundations of mathematics. Hilbert transforms arose during his study of integral equations (Hilbert, D., 1912: *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen.* Teubner, p. 75). (Portrait courtesy of Photo AKG, London, with permission.)

Because of the singularity at $\tau = t$, the integrals in Equation 14.1.4 and Equation 14.1.5 must be taken in the *Cauchy principal value* sense by approaching the singularity point from both sides, namely

$$\int_{-\infty}^{\infty} f(\tau)\,d\tau = \lim_{\epsilon\to 0}\left[\int_{-\infty}^{t-\epsilon} f(\tau)\,d\tau + \int_{t+\epsilon}^{\infty} f(\tau)\,d\tau\right], \tag{14.1.6}$$

so that the infinities to the right and left of $\tau = t$ cancel each other. See Section 10.10. We also note that the Hilbert transform is basically a convolution and does not produce a change of domain; if x is a function of time, then $\widehat{x}$ is also a function of time. This is quite different from what we encountered with Laplace or Fourier transforms.

From its origin in phase shifting, Hilbert transforms of sinusoidal functions are trivial. Some examples are

$$\cos(\widehat{\omega t + \varphi}) = \cos\left(\omega t + \varphi - \tfrac{\pi}{2}\right) = \operatorname{sgn}(\omega)\sin(\omega t + \varphi). \tag{14.1.7}$$

Similarly,

$$\sin(\widehat{\omega t + \varphi}) = -\operatorname{sgn}(\omega)\cos(\omega t + \varphi), \tag{14.1.8}$$

and

$$\widehat{e^{i\omega t+i\varphi}} = -i\ \mathrm{sgn}(\omega) e^{i\omega t+i\varphi}. \tag{14.1.9}$$

Thus, Hilbert transformation does not change the amplitude of sine or cosine but does change their phase by $\pm\pi/2$.

• **Example 14.1.1**

Let us apply the integral definition of the Hilbert transform, Equation 14.1.4, to find the Hilbert transform of $\sin(\omega t)$, $\omega \neq 0$.

From the definition,

$$\mathcal{H}\left[\sin(\omega t)\right] = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin(\omega\tau)}{t-\tau}\, d\tau. \tag{14.1.10}$$

If $x = t - \tau$, then

$$\mathcal{H}\left[\sin(\omega t)\right] = -\frac{\cos(\omega t)}{\pi}\int_{-\infty}^{\infty} \frac{\sin(\omega x)}{x}\, dx = -\cos(\omega t)\ \mathrm{sgn}(\omega). \tag{14.1.11}$$

□

• **Example 14.1.2**

Let us compute the Hilbert transform of $x(t) = \sin(t)/(t^2+1)$ from the definition of the Hilbert transform, Equation 14.1.4.

From the definition,

$$\widehat{x}(t) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{\sin(\tau)}{(t-\tau)(\tau^2+1)}\, d\tau = \frac{1}{\pi}\Im\left[PV \int_{-\infty}^{\infty} \frac{e^{i\tau}}{(t-\tau)(\tau^2+1)}\, d\tau\right]. \tag{14.1.12}$$

Because of the singularity on the real axis at $\tau = t$, we treat the integrals in Equation 14.1.12 in the sense of Cauchy principal value.

To evaluate Equation 14.1.12, we convert it into a closed contour integration by introducing a semicircle C_R of infinite radius in the upper half-plane. This yields a closed contour C, which consists of the real line plus this semicircle. Therefore, Equation 14.1.12 can be rewritten

$$PV \int_{-\infty}^{\infty} \frac{e^{i\tau}}{(t-\tau)(\tau^2+1)}\, d\tau = PV \oint_C \frac{e^{iz}}{(t-z)(z^2+1)}\, dz - \int_{C_R} \frac{e^{iz}}{(t-z)(z^2+1)}\, dz. \tag{14.1.13}$$

The second integral on the right side of Equation 14.1.13 vanishes by Equation 10.9.7.

The evaluation of the closed integral in Equation 14.1.13 follows from the residue theorem. We have that

$$\mathrm{Res}\left[\frac{e^{iz}}{(t-z)(z^2+1)}; t\right] = \lim_{z\to t} \frac{(z-t)\, e^{iz}}{(t-z)(z^2+1)} = -\frac{e^{it}}{t^2+1}, \tag{14.1.14}$$

and

$$\mathrm{Res}\left[\frac{e^{iz}}{(t-z)(z^2+1)}; i\right] = \lim_{z\to i} \frac{(z-i)\, e^{iz}}{(t-z)(z^2+1)} = \frac{e^{-1}}{2i(t-i)}. \tag{14.1.15}$$

We do not have a contribution from $z = -i$ because it lies *outside* of the closed contour.

The Hilbert Transform of Some Common Functions

	function, $\mathbf{x}(\mathbf{t})$	Hilbert transform, $\widehat{\mathbf{x}}(\mathbf{t})$
1.	$\begin{cases} 1, & a < t < b \\ 0, & \text{otherwise} \end{cases}$	$\dfrac{1}{\pi} \ln \left\lvert \dfrac{t-a}{t-b} \right\rvert$
2.	$\sin(\omega t + \varphi)$	$-\operatorname{sgn}(\omega) \cos(\omega t + \varphi)$
3.	$\cos(\omega t + \varphi)$	$\operatorname{sgn}(\omega) \sin(\omega t + \varphi)$
4.	$e^{i\omega t + \varphi i}$	$-i \ \operatorname{sgn}(\omega) e^{i\omega t + \varphi i}$
5.	$\dfrac{1}{t}$	$-\pi \delta(t)$
6.	$\dfrac{1}{t^2 + a^2}, \quad 0 < \Re(a)$	$\dfrac{t}{a(t^2 + a^2)}$
7.	$\dfrac{\lambda t + \mu a}{t^2 + a^2}, \quad 0 < \Re(a)$	$\dfrac{\mu t - \lambda a}{t^2 + a^2}$
8.	$\dfrac{1}{1 + t^4}$	$\dfrac{t(1 + t^2)}{\sqrt{2}\,(1 + t^4)}$
9.	$\dfrac{\sin(at)}{t}, \quad 0 < a$	$\dfrac{1 - \cos(at)}{t}$
10.	$\dfrac{\sin(t)}{1 + t^2}$	$\dfrac{e^{-1} - \cos(t)}{1 + t^2}$
11.	$\sin(at) J_1(at), \quad 0 < a$	$-\cos(at) J_1(at)$
12.	$\sin(at) J_n(bt), \quad 0 < b < a$	$-\cos(at) J_n(bt)$
13.	$\cos(at) J_1(at), \quad 0 < a$	$\sin(at) J_1(at)$
14.	$\cos(at) J_n(bt), \quad 0 < b < a$	$\sin(at) J_n(at)$
15.	$\begin{cases} \sqrt{a^2 - t^2}, & -a < t < a \\ 0, & \text{otherwise} \end{cases}$	$\begin{cases} t + \sqrt{t^2 - a^2}, & -\infty < t < -a \\ t, & -a < t < a \\ t - \sqrt{t^2 - a^2}, & a < t < \infty \end{cases}$
16.	$\sin\left(a\sqrt{t}\,\right) H(t), \quad 0 < a$	$\begin{cases} -e^{-a\sqrt{\lvert t \rvert}}, & -\infty < t < 0 \\ -\cos\left(a\sqrt{t}\,\right), & 0 < t < \infty \end{cases}$

Therefore,

$$PV \int_{-\infty}^{\infty} \frac{e^{i\tau}}{(t-\tau)(\tau^2+1)}\, d\tau = -\frac{\pi i\, e^{it}}{t^2+1} + \frac{\pi\, e^{-1}(t+i)}{t^2+1}. \qquad \textbf{(14.1.16)}$$

Only one half of the value of the residue at $z = t$ was included; this reflects the semicircular indentation around the singularity there. Substituting Equation 14.1.16 into Equation 14.1.12, we obtain the final result that

$$\mathcal{H}\left[\frac{\sin(t)}{t^2+1}\right] = \frac{e^{-1} - \cos(t)}{t^2+1}. \qquad \textbf{(14.1.17)}$$

□

• Example 14.1.3

Let us employ the relationship that the Fourier transform of $\widehat{x}(t)$ equals $-i\,\mathrm{sgn}(\omega)$ times the Fourier transform of $x(t)$ to find the Hilbert transform of $x(t) = e^{-t^2}$.

Because $\mathcal{F}(e^{-t^2}) = \sqrt{\pi}e^{-\omega^2/4}$,

$$\widehat{X}(\omega) = -i\sqrt{\pi}\,\mathrm{sgn}(\omega)e^{-\omega^2/4}. \qquad \textbf{(14.1.18)}$$

Therefore,

$$\widehat{x}(t) = \frac{i}{2\sqrt{\pi}} \int_{-\infty}^{0} e^{it\omega-\omega^2/4}\, d\omega - \frac{i}{2\sqrt{\pi}} \int_{0}^{\infty} e^{it\omega-\omega^2/4}\, d\omega \qquad \textbf{(14.1.19)}$$

$$= \frac{i}{\sqrt{\pi}} \int_{-\infty}^{0} e^{2it\eta-\eta^2}\, d\eta - \frac{i}{\sqrt{\pi}} \int_{0}^{\infty} e^{2it\eta-\eta^2}\, d\eta \qquad \textbf{(14.1.20)}$$

$$= \frac{e^{-t^2}}{\sqrt{\pi}} \int_{-i\infty}^{t} e^{-s^2}\, ds - \frac{e^{-t^2}}{\sqrt{\pi}} \int_{t}^{i\infty} e^{-s^2}\, ds = \frac{2e^{-t^2}}{\sqrt{\pi}} \int_{0}^{t} e^{-s^2}\, ds, \qquad \textbf{(14.1.21)}$$

where $s = t + \eta i$. The integral in Equation 14.1.21 is the well-known *Dawson's integral*.[3] See Gautschi and Waldvogel[4] for an alternative derivation. □

• Example 14.1.4: Numerical computation of the Hilbert transform

Recently André Weideman[5] devised a particularly efficient method for *numerically* computing the Hilbert transform when $x(t)$ is known exactly for any real t and enjoys the property that

$$\int_{-\infty}^{\infty} |x(t)|^2\, dt < \infty. \qquad \textbf{(14.1.22)}$$

---

[3] Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in Fortran: The Art of Scientific Computing.* Cambridge University Press, Section 6.10.

[4] Gautschi, W., and J. Waldvogel, 2000: Computing the Hilbert transform of the generalized Laguerre and Hermite weight functions. *BIT*, **41**, 490–503.

[5] Weideman, J. A. C., 1995: Computing the Hilbert transform on the real line. *Math. Comput.*, **64**, 745–762.

Given Equation 14.1.22, the function $x(t)$ can be represented by the rational expansion

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \rho_n(t), \tag{14.1.23}$$

where $\rho_n(t)$ is the set of rational functions

$$\rho_n(t) = \frac{(1+it)^n}{(1-it)^{n+1}}, \qquad n = 0, \pm 1, \pm 2, \cdots, \tag{14.1.24}$$

and

$$a_n = \frac{1}{\pi}\int_{-\infty}^{\infty} x(t)\rho_n^*(t)\,dt \tag{14.1.25}$$

or

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[1 - i\tan\left(\tfrac{1}{2}\theta\right)\right] x\left[\tan\left(\tfrac{1}{2}\theta\right)\right] e^{-in\theta}\,d\theta, \tag{14.1.26}$$

if we introduce the substitution $t = \tan(\theta/2)$.

Why is Equation 14.1.23 useful? Taking the Hilbert transform of both sides of this equation,

$$\widehat{x}(t) = \sum_{n=-\infty}^{\infty} a_n \widehat{\rho}_n(t). \tag{14.1.27}$$

Using contour integration, we find that

$$\widehat{\rho}_n(t) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{(1+i\tau)^n}{(1-i\tau)^{n+1}(t-\tau)}\,d\tau = -i\ \mathrm{sgn}(n)\rho_n(t), \tag{14.1.28}$$

where $\mathrm{sgn}(t)$ is the signum function with $\mathrm{sgn}(0) = 1$. Therefore,

$$\widehat{x}(t) = -i \sum_{n=-\infty}^{\infty} \mathrm{sgn}(n)\, a_n\, \rho_n(t). \tag{14.1.29}$$

We must now approximate Equation 14.1.29 so that we can evaluate it numerically. We do this by introducing the following truncated version:

$$\widehat{x}_N(t) = -i \sum_{n=-N}^{N-1} \mathrm{sgn}(n)\, A_n\, \rho_n(t). \tag{14.1.30}$$

This particular truncation was chosen because $\rho_n(t)$ and $\rho_{-n-1}(t)$ are a conjugate pair. The coefficient a_n has become A_n, which equals

$$A_n = \frac{1}{N} \sum_{j=-N+1}^{N-1} \left[1 - i\tan\left(\tfrac{1}{2}\theta_j\right)\right] x\left[\tan\left(\tfrac{1}{2}\theta_j\right)\right] e^{-in\theta_j}, \tag{14.1.31}$$

where $\theta_j = \pi j/N$. The terms corresponding to $j = \pm N$ have been set to zero because it is assumed that $x(t)$ vanishes rapidly with $t \to \pm\infty$. Finally, we substitute θ for t and transform Equation 14.1.30 into

$$\widehat{x}_N(t_j) = -\frac{i}{1 - i\tan(\theta_j)} \sum_{n=-N}^{N-1} \mathrm{sgn}(n) A_n e^{in\theta_j}. \tag{14.1.32}$$

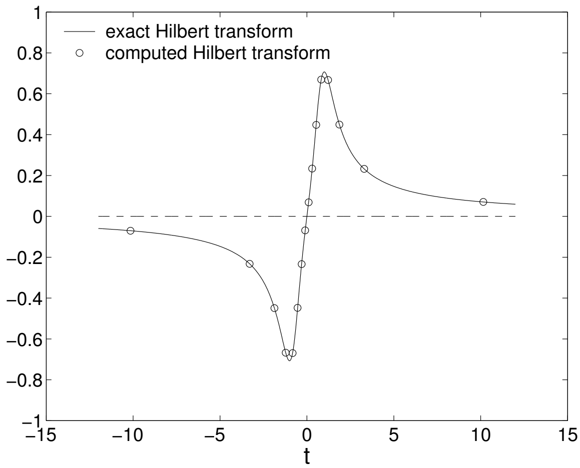


Figure 14.1.1: The Hilbert transform for $x(t) = 1/(1+t^4)$ computed from Weideman's algorithm.

The advantage of Equation 14.1.31 and Equation 14.1.32 is that they can be evaluated using fast Fourier transforms. For example, the following MATLAB script devised by Weideman illustrates his methods for $x(t) = 1/(1+t^4)$:

```
% initialize parameters used in computation
b = 1; N = 8; n = [-N:N-1]';
% set up collocation points and evaluate function there
t = b * tan(pi*(n+1/2)/(2*N)); F = 1./(1+t.^4);
% evaluate Equation 14.1.31
an = fftshift(fft(F.*(b-i*t)));
% compute Hilbert transform via Equation 14.1.32
hilbert = ifft(fftshift(i*(sign(n+1/2).*an)))./(b-i*t);
hilbert = -real(hilbert);
% find points at which we will compute exact answer
tt = [-12:0.02:12];
% compute exact answer
answer = tt.*(1+tt.^2)./(1+tt.^4)./sqrt(2);
fzero = zeros(size(tt));
% plot both computed Hilbert transform and exact answer
plot(tt,answer,'-',t,hilbert,'o',tt,fzero,'--')
xlabel('t','Fontsize',20)
legend('exact Hilbert transform','computed Hilbert transform')
legend boxoff
```

Figure 14.1.1 illustrates Weideman's algorithm for numerically computing the Hilbert transform of $1/(1+t^4)$.

There are two important points concerning Weideman's implementation of his algorithm. First, the collocation points originally given by $t_j = \tan[\pi j/(2N)]$, $j = -N, \ldots, N-1$ have changed to $t_j = \tan[(j+\frac{1}{2})\pi/(2N)]$, $j = -N, \ldots, N-1$. This change replaces the trapezoidal rule discretization for the Fourier coefficients with a midpoint rule. The advantages are twofold: First, it avoids the nuisance of dealing with a collocation point at infinity. Second, it actually yields more accurate results in many cases.

The discerning student will also notice that Weideman introduced a free parameter b, which we set to one. This rescaling parameter can have a major influence on the accuracy. The interested student is referred to the bottom of page 756 in Weideman's paper for further details. □

• Example 14.1.5: Discrete Hilbert transform

Quite often the function is given as discrete data points. How do we find the Hilbert transform in this case? We will now prove[6] that the equivalent *discrete* Hilbert transform is

$$\mathcal{H}(f_n) = \widehat{f}_k = \begin{cases} \dfrac{2}{\pi} \displaystyle\sum_{n \text{ odd}} \dfrac{f_n}{k-n}, & k \text{ even}, \\ \dfrac{2}{\pi} \displaystyle\sum_{n \text{ even}} \dfrac{f_n}{k-n}, & k \text{ odd}, \end{cases} \tag{14.1.33}$$

where f_n denotes a set of discrete data values that are sampled at $t = nT$ and both k and n run from $-\infty$ to ∞. The corresponding inverse is

$$f_n = \begin{cases} \dfrac{2}{\pi} \displaystyle\sum_{k \text{ odd}} \dfrac{\widehat{f}_k}{k-n}, & n \text{ even}, \\ \dfrac{2}{\pi} \displaystyle\sum_{k \text{ even}} \dfrac{\widehat{f}_k}{k-n}, & n \text{ odd}. \end{cases} \tag{14.1.34}$$

We begin our proof by inserting Equation 14.1.33 into Equation 14.1.34. For n even,

$$f_n = \frac{2}{\pi} \sum_{k \text{ odd}} \frac{1}{k-n} \left(\frac{2}{\pi} \sum_{p \text{ even}} \frac{f_p}{k-p} \right) = \frac{4}{\pi^2} \sum_{p\,\text{even}} \sum_{k\,\text{odd}} \frac{f_p}{(k-p)(k-n)} \tag{14.1.35}$$

$$= \frac{4}{\pi^2} \sum_{k \text{ odd}} \frac{f_n}{(k-n)^2} + \frac{4}{\pi^2} \sum_{p\,\text{even}, p \neq n} \sum_{k\,\text{odd}} (n-p) f_p \left\{ \frac{1}{k-n} - \frac{1}{k-p} \right\}. \tag{14.1.36}$$

The term within the curly brackets equals zero as k runs through all of its values. Therefore, Equation 14.1.36 reduces to

$$f_n = \frac{8}{\pi^2} f_n \left(1 + \frac{1}{3^2} + \frac{1}{5^2} + + \frac{1}{7^2} + \cdots \right). \tag{14.1.37}$$

However, the term in the brackets of Equation 14.1.37 equals $\pi^2/8$. Therefore, Equation 14.1.33 and Equation 14.1.34 is proved for n even. An identical proof follows for n odd.

A popular alternative[7] to Equation 14.1.33 involves the (fast) Fourier transform and the relationship that $\widehat{X}(\omega) = -i\ \mathrm{sgn}(\omega) X(\omega)$, where $X(\omega)$ and $\widehat{X}(\omega)$ denote the Fourier transform of $x(t)$ and $\widehat{x}(t)$, respectively. In this technique, a fast Fourier transform is taken of the data. This transformed dataset is then multiplied by $-i\ \mathrm{sgn}(\omega)$ and then back transformed to give the Hilbert transform.

---

[6] See Kak, S. C., 1970: The discrete Hilbert transform. *Proc. IEEE*, **58**, 585–586. For an alternative derivation, see Kress, R., and E. Martensen, 1970: Anwendung der Rechteckregel auf die reelle Hilbert-transformation mit unendlichem Intervall. *Z. Angew. Math. Mech.*, **50**, T61–T64.

[7] Čížek, V., 1970: Discrete Hilbert transform. *IEEE Trans. Audio Electroacoust.*, **AU-18**, 340–343.

Let $x(t)$ be a real, even function. Then $X(\omega)$, the Fourier transform of $x(t)$, is also an even function. Consequently,

$$\widehat{x}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{X}(\omega)e^{i\omega t}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} -i\,\mathrm{sgn}(\omega)X(\omega)\left[\cos(\omega t) + i\sin(\omega t)\right]\,d\omega \quad \textbf{(14.1.38)}$$

$$= -\frac{i}{2\pi}\int_{-\infty}^{\infty} \mathrm{sgn}(\omega)X(\omega)\cos(\omega t)\,d\omega + \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{sgn}(\omega)X(\omega)\sin(\omega t)\,d\omega \quad \textbf{(14.1.39)}$$

$$= \frac{1}{\pi}\int_0^{\infty} X(\omega)\sin(\omega t)\,d\omega. \quad \textbf{(14.1.40)}$$

Note that the Hilbert transform in this case is an odd function. Similarly, if $x(t)$ is a real, odd function,

$$\widehat{x}(t) = -\frac{i}{\pi}\int_0^{\infty} X(\omega)\cos(\omega t)\,d\omega, \quad \textbf{(14.1.41)}$$

and the Hilbert transform is an even function.

Problems

1. Show that the Hilbert transform of a constant function is zero.

2. Use Equation 14.1.4 to compute the Hilbert transform of $\cos(\omega t)$, $\omega \neq 0$.

3. Use Equation 14.1.4 to show that the Hilbert transform of the Dirac delta function $\delta(t)$ is $1/(\pi t)$.

4. Use Equation 14.1.4 to show that the Hilbert transform of $1/(t^2+1)$ is $t/(t^2+1)$.

5. The output $y(t)$ from an ideal lowpass filter can be expressed by the convolution integral

$$y(t) = x(t) * \frac{\sin(2\pi\omega t)}{\pi t},$$

where $x(t)$ is the input signal. Show that this expression can also be expressed in terms of Hilbert transforms as

$$y(t) = \mathcal{H}[x(t)\cos(2\pi\omega t)]\sin(2\pi\omega t) - \mathcal{H}[x(t)\sin(2\pi\omega t)]\cos(2\pi\omega t).$$

Following Example 14.1.3, find the Hilbert transforms of

6. $x(t) = \dfrac{1}{1+t^2}$

7. $x(t) = \begin{cases} 1, & -a < t < a \\ 0, & \text{otherwise} \end{cases}$

8. Using the commutative and associate properties of convolution, $f(t) * g(t) = g(t) * f(t)$ and $[f(t) * g(t)] * v(t) = f(t) * [g(t) * v(t)]$, respectively, and the definition of the Hilbert transform, Equation 14.1.4, show[8] that

$$\mathcal{H}[f(t) * g(t)] = \widehat{f}(t) * g(t) = f(t) * \widehat{g}(t).$$

[8] For an application, see Sakai, H., and G. A. Vanasse, 1966: Hilbert transform in Fourier spectroscopy. *J. Opt. Soc. Am.*, **56**, 131–132.

Using MATLAB, test Weideman's algorithm for the following cases. Why does the algorithm do well or not?

9. $\begin{cases} 1, & -1 < t < 1 \\ 0, & \text{otherwise} \end{cases}$ 10. $\sin(t)$ 11. $\dfrac{1}{t^2+1}$ 12. $\dfrac{\sin(t)}{1+t^4}$

For Problem 12, you will need

$$\mathcal{H}\left[\frac{\sin(t)}{t^4+1}\right] = \frac{e^{-1/\sqrt{2}}[\cos(1/\sqrt{2}\,) + \sin(1/\sqrt{2}\,)t^2] - \cos(t)}{t^4+1}.$$

14.2 SOME USEFUL PROPERTIES

In principle we could construct any desired transform from the definition of the Hilbert transform. However, there are several general theorems that are much more effective in finding new transforms.

Linearity

From the definition of the Hilbert transform, it immediately follows that if $z(t) = c_1x(t) + c_2y(t)$, where c_1 and c_2 are arbitrary constants, then $\widehat{z}(t) = c_1\widehat{x}(t) + c_2\widehat{y}(t)$.

The energy in a signal and its Hilbert transform are the same.

Consider the energy spectral densities at input and output of a quadrature phase shifter. The output equals

$$|\widehat{X}(\omega)|^2 = \left|\mathcal{F}[\widehat{x}(t)]\right|^2 = |-i\ \mathrm{sgn}(\omega)|^2|X(\omega)|^2 = |X(\omega)|^2. \qquad \textbf{(14.2.1)}$$

Because the energy spectral density at input and output are the same, so are the total energies.

A signal and its Hilbert transform are orthogonal.

From Parseval's theorem,

$$\int_{-\infty}^{\infty} x(t)\widehat{x}(t)\,dt = \int_{-\infty}^{\infty} X(\omega)\widehat{X}^*(\omega)\,d\omega, \qquad \textbf{(14.2.2)}$$

where $\widehat{X}(\omega) = \mathcal{F}[\widehat{x}(t)]$. Then,

$$\int_{-\infty}^{\infty} X(\omega)\widehat{X}^*(\omega)\,d\omega = \int_{-\infty}^{\infty} i\ \mathrm{sgn}(\omega)|X(\omega)|^2\,d\omega = 0, \qquad \textbf{(14.2.3)}$$

because the integrand in the middle expression of Equation 14.2.3 is odd. Thus,

$$\int_{-\infty}^{\infty} x(t)\widehat{x}(t)\,dt = 0. \qquad \textbf{(14.2.4)}$$

The reason why a function and its Hilbert transform are orthogonal to each other follows from the fact that a Hilbert transformation of a function shifts the phase of each Fourier component of the function *forward* by $\pi/2$ for positive frequencies and *backward* for negative frequencies.

• Example 14.2.1

Let us verify the orthogonality condition for Hilbert transforms using $x(t) = 1/(1+t^2)$. Because $\widehat{x}(t) = t/(1+t^2)$,

$$\int_{-\infty}^{\infty} x(t)\widehat{x}(t)\, dt = \int_{-\infty}^{\infty} \frac{t}{(1+t^2)^2}\, dt = 0, \tag{14.2.5}$$

since the integrand is an odd function. □

Shifting

Let us find the Hilbert transform of $x(t+a)$ if we know $\widehat{x}(t)$. From the definition of Hilbert transforms,

$$\mathcal{H}[x(t+a)] = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\eta+a)}{t-\eta}\, d\eta = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\tau)}{(t+a)-\tau}\, d\tau = \widehat{x}(t+a) \tag{14.2.6}$$

or $\mathcal{H}[x(t+a)] = \widehat{x}(t+a)$.

Time scaling

Let $a > 0$. Then,

$$\mathcal{H}[x(at)] = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(a\eta)}{t-\eta}\, d\eta = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\tau)}{at-\tau}\, d\tau = \widehat{x}(at). \tag{14.2.7}$$

On the other hand, if $a < 0$,

$$\mathcal{H}[x(at)] = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(a\eta)}{t-\eta}\, d\eta = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(\tau)}{at-\tau}\, d\tau = -\widehat{x}(at). \tag{14.2.8}$$

Thus, we have that $\mathcal{H}[x(at)] = \operatorname{sgn}(a)\, \widehat{x}(at)$.

Derivatives

Let us find the relationship between the nth derivative of $x(t)$ and its Hilbert transform. Using the derivative rule as it applies to Fourier transforms,

$$\mathcal{H}\left\{\mathcal{F}\left[\frac{d^n x}{dt^n}\right]\right\} = -i\operatorname{sgn}(\omega)(i\omega)^n X(\omega) = (i\omega)^n[-i\operatorname{sgn}(\omega)X(\omega)] = (i\omega)^n \widehat{X}(\omega) = \mathcal{F}\left[\frac{d^n \widehat{x}}{dt^n}\right]. \tag{14.2.9}$$

Taking the inverse Fourier transforms, we have that

$$\mathcal{H}\left(\frac{d^n x}{dt^n}\right) = \frac{d^n \widehat{x}}{dt^n}. \tag{14.2.10}$$

Some General Properties of Hilbert Transforms

	function, $\mathbf{x(t)}$	Hilbert transform, $\widehat{\mathbf{x}}(\mathbf{t})$
1.	$\widehat{x}(t)$	$-x(t)$
2.	$x(t)+y(t)$	$\widehat{x}(t)+\widehat{y}(t)$
3.	$x(t+a), \quad a$ real	$\widehat{x}(t+a)$
4.	$\dfrac{d^n x(t)}{dt^n}$	$\dfrac{d^n \widehat{x}(t)}{dt^n}$
5.	$x(at)$	$\operatorname{sgn}(a)\,\widehat{x}(at)$
6.	$tx(t)$	$t\widehat{x}(t)+\frac{1}{\pi}\int_{-\infty}^{\infty} x(\tau)\,d\tau$
7.	$(t+a)x(t)$	$(t+a)\widehat{x}(t)+\frac{1}{\pi}\int_{-\infty}^{\infty} x(\tau)\,d\tau$

Convolution

Hilbert transforms enjoy a similar, but not identical, property with Fourier transforms with respect to convolution. If

$$w(t) = u(t) * v(t) = \int_{-\infty}^{\infty} u(\tau)v(t-\tau)\,d\tau = \int_{-\infty}^{\infty} u(t-\tau)v(\tau)\,d\tau, \qquad \textbf{(14.2.11)}$$

then

$$\widehat{w}(t) = v(t) * \widehat{u}(t). \qquad \textbf{(14.2.12)}$$

Proof: From the convolution theorem for Fourier transforms, $W(\omega) = V(\omega)U(\omega)$. Multiplying both sides of the equation by $-i\operatorname{sgn}(\omega)$,

$$\widehat{W}(\omega) = -i\operatorname{sgn}(\omega)W(\omega) = V(\omega)[-i\operatorname{sgn}(\omega)U(\omega)] = V(\omega)\widehat{U}(\omega). \qquad \textbf{(14.2.13)}$$

Again, using the convolution theorem as it applies to Fourier transforms, we arrive at the final result. □

• **Example 14.2.2**

Given the functions $u(t) = \cos(t)$ and $v(t) = 1/(1+t^4)$, let us verify the convolution theorem as it applies to Hilbert transforms.

With $u(t) = \cos(t)$ and $v(t) = 1/(1+t^4)$,

$$w(t) = u(t) * v(t) = \int_{-\infty}^{\infty} \frac{\cos(t-x)}{1+x^4}\,dx \tag{14.2.14}$$

$$= \int_{-\infty}^{\infty} \frac{\cos(t)\cos(x)}{1+x^4}\,dx + \int_{-\infty}^{\infty} \frac{\sin(t)\sin(x)}{1+x^4}\,dx \tag{14.2.15}$$

$$= \frac{\pi}{\sqrt{2}} e^{-1/\sqrt{2}} \left[\cos\left(\frac{1}{\sqrt{2}}\right) + \sin\left(\frac{1}{\sqrt{2}}\right)\right] \cos(t) \tag{14.2.16}$$

so that

$$\widehat{w}(t) = \frac{\pi}{\sqrt{2}} e^{-1/\sqrt{2}} \left[\cos\left(\frac{1}{\sqrt{2}}\right) + \sin\left(\frac{1}{\sqrt{2}}\right)\right] \sin(t). \tag{14.2.17}$$

Because $\widehat{v}(t) = t(1+t^2)/[\sqrt{2}\,(1+t^4)]$,

$$u(t) * \widehat{v}(t) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \cos(t-x) \frac{x(1+x^2)}{1+x^4}\,dx \tag{14.2.18}$$

$$= \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \frac{\cos(t)\cos(x) x(1+x^2)}{1+x^4}\,dx + \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \frac{\sin(t)\sin(x) x(1+x^2)}{1+x^4}\,dx \tag{14.2.19}$$

$$= \frac{1}{\sqrt{2}} \sin(t) \int_{-\infty}^{\infty} \frac{x(1+x^2)\sin(x)}{1+x^4}\,dx \tag{14.2.20}$$

$$= \frac{\pi}{\sqrt{2}} e^{-1/\sqrt{2}} \left[\cos\left(\frac{1}{\sqrt{2}}\right) + \sin\left(\frac{1}{\sqrt{2}}\right)\right] \sin(t), \tag{14.2.21}$$

and the convolution theorem for Hilbert transforms holds true in this case. □

Product theorem

Let $f(t)$ and $g(t)$ denote complex functions with Fourier transforms $F(\omega)$ and $G(\omega)$, respectively. If

1) $F(\omega)$ vanishes for $|\omega| > a$, and $G(\omega)$ vanishes for $|\omega| < a$, where $a > 0$,

or

2) $f(t)$ and $g(t)$ are analytic functions (their real and imaginary parts are Hilbert pairs),

then the Hilbert transform of the product of $f(t)$ and $g(t)$ is

$$\mathcal{H}[f(t)g(t)] = f(t)\widehat{g}(t). \tag{14.2.22}$$

Proof:[9] The product $f(t)g(t)$ can be expressed as

$$f(t)g(t) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u)G(v)e^{i(u+v)t}\,dv\,du. \tag{14.2.23}$$

[9] See Bedrosian, E., 1963: A product theorem for Hilbert transforms. *Proc. IEEE*, **51**, 868–869. This theorem has been extended to functions of n-dimensional real vectors by Stark, H., 1971: An extension of the Hilbert transform product theorem. *Proc. IEEE*, **59**, 1359–1360.

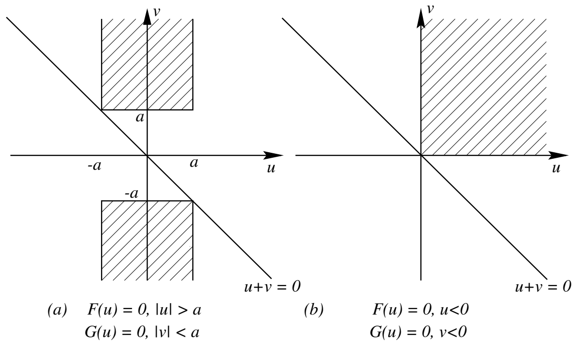


Figure 14.2.1: Region of integration in the proof of the product theorem.

Because $\mathcal{H}(e^{ibt}) = i\ \text{sgn}(b)e^{ibt}$,

$$\mathcal{H}[f(t)g(t)] = \frac{i}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u)G(v)\ \text{sgn}(u+v)e^{i(u+v)t}\,dv\,du. \tag{14.2.24}$$

The shaded regions of Figure 14.2.1 are those in which the product $F(u)G(v)$ is nonvanishing for the conditions of the theorem. In Figure 14.2.1(a) the nonoverlapping Fourier transforms yield two semi-infinite strips in which the product is nonvanishing. In Figure 14.2.1(b), for analytic functions, the Fourier transforms vanish for negative arguments[10] so that the product is nonvanishing only in the first quadrant. In both cases $\text{sgn}(u+v) = \text{sgn}(v)$ over the regions of integration in which the integrand is nonvanishing. Thus,

$$\mathcal{H}[f(t)g(t)] = \frac{i}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u)G(v)\ \text{sgn}(v)e^{i(u+v)t}\,dv\,du \tag{14.2.25}$$

$$= f(t)\frac{i}{2\pi}\int_{-\infty}^{\infty} G(v)\ \text{sgn}(v)e^{ivt}\,dv = f(t)\widehat{g}(t). \tag{14.2.26}$$

□

• Example 14.2.3: Hilbert transforms of band-pass functions

In communications, we have the double-sideband, amplitude-modulated signal given by $a(t)\cos(\omega t + \varphi)$, where φ is constant. From the product theorem its Hilbert transform equals $a(t)\sin(\omega t + \varphi)$, $\omega > 0$, provided that the highest frequency component in $a(t)$ is less than ω. Paradoxically, the Hilbert transform of more general $a(t)\cos[\omega t + \varphi(t)]$, which equals $a(t)\sin[\omega t + \varphi(t)]$, has no such restriction.

Problems

Verify the orthogonality property of Hilbert transforms using

1. $x(t) = 1/(1 + t^4)$

2. $x(t) = \sin(t)/(1 + t^2)$

3. $x(t) = \begin{cases} 1, & 0 < t < a \\ 0, & \text{otherwise} \end{cases}$

---

[10] Titchmarsh, E. C., 1948: *Introduction to the Theory of Fourier Integrals*. Oxford University Press, p. 128.

Verify the convolution theorem for Hilbert transforms using

4. $u(t) = \begin{cases} 1, & 0 < t < a, \\ 0, & \text{otherwise,} \end{cases} \quad v(t) = \sin(t)$

5. $u(t) = \cos(t), \quad v(t) = \dfrac{1}{1+t^2}$

6. Use the product theorem to show that

$$\mathcal{H}[\sin(at)J_n(bt)] = -\cos(at)J_n(bt), \qquad 0 < b < a,$$

if $n = 0, 1, 2, 3, \ldots$.
Hint:

$$\mathcal{F}[J_n(bt)] = \frac{2(-1)^m}{\sqrt{b^2 - \omega^2}} T_n\left(\frac{|\omega|}{b}\right) H(b - |\omega|),$$

where $T_n(\cdot)$ is a Chebyshev polynomial of the first kind and $m = n/2$ or $(n-1)/2$, depending upon which definition gives an integer.

7. Given cosine and sine integrals:

$$Ci(x) = -\int_x^\infty \frac{\cos(t)}{t}\,dt, \qquad Si(x) = -\int_x^\infty \frac{\sin(t)}{t}\,dt,$$

and

$$\mathcal{H}[Ci(a|t|)] = -\operatorname{sgn}(t)Si(a|t|), \quad 0 < a,$$

use the product rule to show that

$$\mathcal{H}[\sin(bt)Ci(a|t|)] = -\operatorname{sgn}(t)\sin(bt)Si(a|t|), \qquad 0 < b < a.$$

Hint:

$$\mathcal{F}[Ci(a|t|)] = \begin{cases} 0, & 0 < |\omega| < a, \\ -\pi/|\omega|, & a < |\omega| < \infty, \end{cases} \qquad 0 < a.$$

8. Prove that

$$\mathcal{H}[tx(t)] = t\widehat{x}(t) - \frac{1}{\pi}\int_{-\infty}^\infty x(\tau)\,d\tau.$$

Hint:

$$\frac{\tau x(\tau)}{t-\tau} = \frac{tx(\tau)}{t-\tau} - x(\tau).$$

14.3 ANALYTIC SIGNALS

The monochromatic signal $A\cos(\omega_0 t + \varphi)$ appears in many physical and engineering applications. It is common to represent this signal by the complex representation $Ae^{i(\omega_0 t+\varphi)}$. These two representations are related to each other by

$$A\cos(\omega_0 t + \varphi) = \Re\left[Ae^{i(\omega_0 t+\varphi)}\right] = \tfrac{1}{2}\left[Ae^{i(\omega_0 t+\varphi)} + Ae^{-i(\omega_0 t+\varphi)}\right]. \qquad \textbf{(14.3.1)}$$

Furthermore, the Fourier transform of $A\cos(\omega_0 t + \varphi)$ is

$$\mathcal{F}[A\cos(\omega_0 t + \varphi)] = \tfrac{1}{2}\left[Ae^{i\varphi}\delta(\omega - \omega_0) + Ae^{-i\varphi}\delta(\omega + \omega_0)\right], \qquad \textbf{(14.3.2)}$$

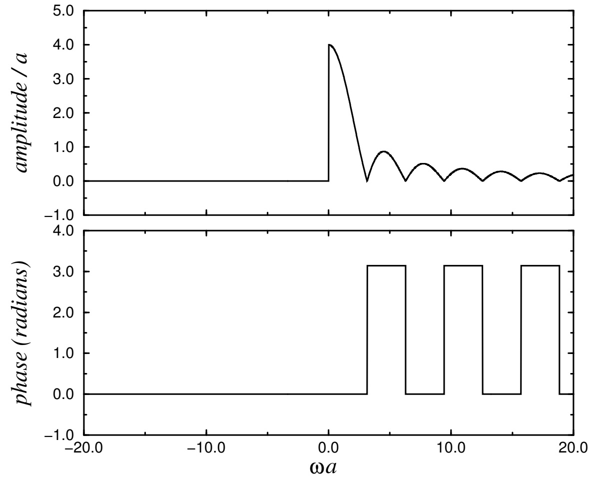


Figure 14.3.1: The spectrum of the analytic signal when $x(t)$ is the rectangular pulse given by Equation 11.1.9.

while the Fourier transform of $Ae^{i(\omega_0 t+\varphi)}$ is

$$\mathcal{F}\left[Ae^{i(\omega_0 t+\varphi)}\right] = Ae^{i\varphi}\delta(\omega-\omega_0). \tag{14.3.3}$$

As Equation 14.3.2 and Equation 14.3.3 clearly show, in passing from the real signal to its complex representation, we double the strength of the positive frequencies and remove entirely the negative frequencies.

Let us generalize these concepts to nonmonochromatic signals. For the real signal $x(t)$ with Fourier transform $X(\omega)$ and the complex signal $z(t)$ with Fourier transform $Z(\omega)$, the previous paragraph shows that our generalization must have the property:

$$Z(\omega) = X(\omega) + \operatorname{sgn}(\omega)X(\omega) \tag{14.3.4}$$

or

$$Z(\omega) = \begin{cases} 2X(\omega), & \omega > 0, \\ X(\omega), & \omega = 0, \\ 0, & \omega < 0. \end{cases} \tag{14.3.5}$$

Taking the inverse of Equation 14.3.4, we have the definition of an *analytic signal* as

$$z(t) = x(t) + i\widehat{x}(t), \tag{14.3.6}$$

where $x(t)$ is a real signal and $\widehat{x}(t)$ is its Hilbert transform.

• Example 14.3.1

In Figure 14.3.1 the amplitude spectrum of the analytic signal is graphed when $x(t)$ is the rectangular pulse, Equation 11.1.9. Note that the amplitude spectrum equals zero for $\omega < 0$ and twice the amplitude spectrum for $\omega > 0$.

□

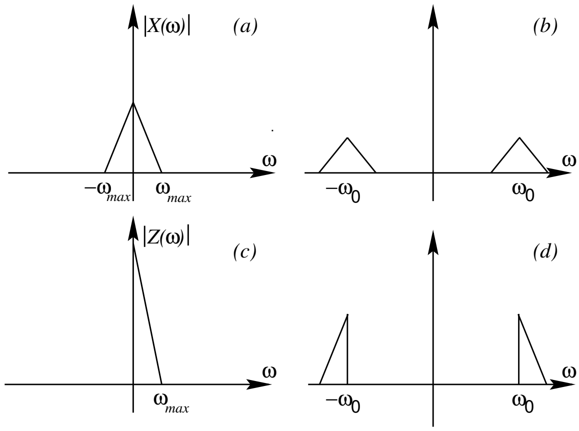


Figure 14.3.2: Given a function $x(t)$ with an amplitude spectrum shown in (a), frame (b) shows the amplitude spectrum of the amplitude-modulated signal $x(t)\cos(\omega_0 t)$ while frames (c) and (d) give the amplitude spectrum of the analytic signal $z(t)$ and $x(t)\cos(\omega_0 t) - \widehat{x}(t)\sin(\omega_0 t)$, respectively.

• Example 14.3.2

Let us find the energy of an analytic signal.
The energy of an analytic signal is

$$\int_{-\infty}^{\infty} |z(t)|^2\, dt = \int_{-\infty}^{\infty} x^2(t)\, dt + \int_{-\infty}^{\infty} \widehat{x}^2(t)\, dt = 2\int_{-\infty}^{\infty} x^2(t)\, dt = 2\int_{-\infty}^{\infty} |X(\omega)|^2\, d\omega \tag{14.3.7}$$

by Parseval's theorem. Thus, the analytic signal has twice the energy of the corresponding real signal. □

Consider the function $x(t)$ whose amplitude spectrum is shown in Figure 14.3.2(a). If we were to amplitude modulate $x(t)$ with $\cos(\omega_0 t)$, then the amplitude spectrum of this modulated signal would appear as pictured in Figure 14.3.2(b).

Consider now the signal

$$y(t) = x(t)\cos(\omega_0 t) - \widehat{x}(t)\sin(\omega_0 t) = \Re\left\{[x(t) + i\widehat{x}(t)]e^{i\omega_0 t}\right\} \tag{14.3.8}$$

$$= \Re\left\{z(t)e^{i\omega_0 t}\right\} = \tfrac{1}{2}\left[z(t)e^{i\omega_0 t} + z^*(t)e^{-i\omega_0 t}\right], \tag{14.3.9}$$

where $z(t)$ is the analytic signal of $x(t)$. We have plotted the amplitude spectrum $|Z(\omega)|$ in Figure 14.3.2(c). If we computed the amplitude spectrum of $y(t)$, we would find that

$$Y(\omega) = \tfrac{1}{2}Z(\omega - \omega_0) + \tfrac{1}{2}Z(-\omega - \omega_0) \tag{14.3.10}$$

$$= \begin{cases} X(\omega - \omega_0), & \omega_0 \le \omega \le \omega_0 + \omega_{\max}, \\ X^*(-\omega - \omega_0), & -\omega_0 - \omega_{\max} \le \omega \le \omega_0, \\ 0, & \text{otherwise.} \end{cases} \tag{14.3.11}$$

We have sketched this amplitude spectrum $|Y(\omega)|$ in Figure 14.3.2(d). Each triangular part is called the *single sideband signal* because it contains the upper frequencies ($|\omega| > \omega_0$) of the modulated signal $x(t)\cos(\omega_0 t)$. Similarly, if we had used $x(t)\cos(\omega_0 t) + \widehat{x}(t)\sin(\omega_0 t)$, we would only have obtained the lower sidebands. Consequently, a communication system using $x(t)\cos(\omega_0 t) - \widehat{x}(t)\sin(\omega_0 t)$ or $x(t)\cos(\omega_0 t) + \widehat{x}(t)\sin(\omega_0 t)$ would realize a 50% savings in its frequency bandwidth over one transmitting $x(t)\cos(\omega_0 t)$.

Problems

1. Find the analytic signal corresponding to $x(t) = \cos(\omega t)$, $\omega > 0$.

2. Show that the polar form of an analytic signal can be written

$$z(t) = |z(t)|e^{i\varphi(t)},$$

where

$$|z(t)|^2 = x^2(t) + \widehat{x}^2(t), \quad \varphi(t) = \tan^{-1}\left[\frac{\widehat{x}(t)}{x(t)}\right].$$

3. Analytic signals are often used with narrow-band waveforms with carrier frequency ω_0. If $\varphi(t) = \omega_0 t + \varphi'(t)$, show that the analytic signal can be written $z(t) = r(t)e^{i\omega_0 t}$, where $r(t) = |z(t)|e^{i\varphi'(t)}$. The function $r(t)$ is called the *complex envelope* or the *phasor amplitude*; this is a generalization of the phasor idea beyond pure alternating currents.

14.4 CAUSALITY: THE KRAMERS-KRONIG RELATIONSHIP

Causality is the physical principle which states that an event cannot proceed its cause. In this section we explore what effect this principle has on Hilbert transforms.

We begin by introducing the concept of causal functions. A *causal function* is a function that equals zero for all $t < 0$. As with all functions we can write it in terms of an even $x_e(t)$ and an odd $x_o(t)$ part as $x(t) = x_e(t) + x_o(t)$. Because $x(t)$ is causal, $x_o(t) = \operatorname{sgn}(t)x_e(t)$ and

$$x(t) = x_e(t) + \operatorname{sgn}(t)x_e(t). \tag{14.4.1}$$

Taking the Fourier transform of Equation 14.4.1, we find that the Fourier transform of *all* causal functions are of the form

$$X(\omega) = X_e(\omega) - i\widehat{X}_e(\omega), \tag{14.4.2}$$

where

$$\widehat{X}_e(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{X_e(\tau)}{\omega - \tau}\,d\tau, \tag{14.4.3}$$

and

$$X_e(\omega) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\widehat{X}_e(\tau)}{\omega - \tau}\,d\tau, \tag{14.4.4}$$

because

$$2\pi\mathcal{F}[x_e(t)\operatorname{sgn}(t)] = \frac{2}{i\omega} * X_e(\omega) = \frac{2}{i}\int_{-\infty}^{\infty} \frac{X_e(\tau)}{\omega - \tau}d\tau. \tag{14.4.5}$$

Equations 14.4.3 and 14.4.4 first arose in dielectric theory and, taken together, are called the *Kramers*[11] *and Kronig*[12] *relation* after their discoverers, who derived these relationships during their work on the dispersion of light by gaseous atoms or molecules.

---

[11] Kramers, H. A., 1929: Die Dispersion und Absorption von Röntgenstrahlen. *Phys. Z.*, **30**, 522–523.

[12] Kronig, R. de L., 1926: On the theory of dispersion of x-rays. *J. Opt. Soc. Am.*, **12**, 547–551.

• **Example 14.4.1**

Let us verify the Kramers-Kronig relation using the causal time function $x(t) = H(t)$.

Because $x_e(t) = \frac{1}{2}$ and $X_e(\omega) = \pi\delta(\omega)$ by Equation 11.2.3,

$$\widehat{X}_e(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\pi\,\delta(\tau)}{\omega - \tau}\,d\tau = -\frac{1}{\omega}. \tag{14.4.6}$$

Consequently, by the Kramers-Kronig relation,

$$\mathcal{F}[H(t)] = X_e(\omega) - i\widehat{X}_e(\omega) = \pi\delta(\omega) + \frac{i}{\omega}. \tag{14.4.7}$$

This agrees with the result given in Example 11.2.2. □

• **Example 14.4.2**

A simple example of a causal function is the impulse response or Green's function introduced in earlier chapters. From Equation 14.4.2 we have the result that the transfer function $G(\omega)$, the Fourier transform of the impulse response, must yield the Hilbert transform pair $G_e(\omega) - i\widehat{G}_e(\omega)$.

For example, if $g(t) = e^{-t}H(t)$, then $G(\omega) = 1/(1 + i\omega)$. Because

$$\frac{1}{1 + i\omega} = \frac{1}{\omega^2 + 1} - i\frac{\omega}{\omega^2 + 1}, \tag{14.4.8}$$

we have the Hilbert transform pair of

$$x(t) = \frac{1}{t^2 + 1} \qquad \text{and} \qquad \widehat{x}(t) = \frac{t}{t^2 + 1}. \tag{14.4.9}$$

□

• **Example 14.4.3**

Let us verify the Kramers-Kronig relation for the Hilbert transform pair

$$x(t) = \frac{1}{t^4 + 1} \qquad \text{and} \qquad \widehat{x}(t) = \frac{t(t^2 + 1)}{\sqrt{2}(t^4 + 1)} \tag{14.4.10}$$

by direct integration.

From Equation 14.4.3, we have that

$$\frac{\omega(\omega^2 + 1)}{\sqrt{2}(\omega^4 + 1)} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{d\tau}{(\tau^4 + 1)(\omega - \tau)}. \tag{14.4.11}$$

Applying the residue theorem to the right side of Equation 14.4.11, we obtain

$$\begin{aligned}\frac{\omega(\omega^2 + 1)}{\sqrt{2}(\omega^4 + 1)} &= i\,\text{Res}\left[\frac{1}{(z^4 + 1)(\omega - z)}; \omega\right] + 2i\,\text{Res}\left[\frac{1}{(z^4 + 1)(\omega - z)}; e^{\pi i/4}\right] \\ &\quad + 2i\,\text{Res}\left[\frac{1}{(z^4 + 1)(\omega - z)}; e^{3\pi i/4}\right].\end{aligned} \tag{14.4.12}$$

We only include one half of the value of the residue at $\tau = \omega$ because the singularity lies on the path of integration and we must treat this integration along the lines of a Cauchy principal value. Evaluating the residues, we find

$$\operatorname{Res}\left[\frac{1}{(z^4+1)(\omega-z)};\omega\right] = -\frac{1}{\omega^4+1}, \tag{14.4.13}$$

$$\operatorname{Res}\left[\frac{1}{(z^4+1)(\omega-z)};e^{\pi i/4}\right] = \frac{\sqrt{2}-(1+i)\omega}{4\sqrt{2}\left[\left(\omega-\frac{1}{\sqrt{2}}\right)^2+\frac{1}{2}\right]}, \tag{14.4.14}$$

and

$$\operatorname{Res}\left[\frac{1}{(z^4+1)(\omega-z)};e^{3\pi i/4}\right] = \frac{\sqrt{2}+(1-i)\omega}{4\sqrt{2}\left[\left(\omega+\frac{1}{\sqrt{2}}\right)^2+\frac{1}{2}\right]}. \tag{14.4.15}$$

Substituting Equation 14.4.13 through Equation 14.4.15 into the right side of Equation 14.4.12, we obtain the left side.

Problems

1. For a causal function $x(t)$, prove that $x_o(t) = \operatorname{sgn}(t)x_e(t)$ and $x_e(t) = \operatorname{sgn}(t)x_o(t)$.

2. Redo our analysis if $x(t)$ is a negative time function, i.e., $x(t) = 0$ if $t > 0$. Verify your result using $x(t) = e^t H(-t)$.

3. Using $g(t) = te^{-t}H(t)$, find the corresponding Hilbert transform pairs.

4. Using $g(t) = e^{-t}\cos(\omega t)H(t)$, find the corresponding Hilbert transform pairs.

5. Verify the Kramers-Kronig relation for the Hilbert transform pair:

$$x(t) = \frac{1}{t^2+1} \qquad \text{and} \qquad \widehat{x}(t) = \frac{t}{t^2+1}$$

by direct integration.

Further Reading

Hahn, S. L., 1996: *Hilbert Transforms in Signal Processing.* Artech House, 442 pp. Covers the basic theory and gives some practical applications.

Chapter 15

Green's Functions

We have devoted a major portion of this book to solving linear ordinary and partial differential solutions. For example, in the case of partial differential equations we introduced the method of separation of variables, which leads to a solution in terms of an eigenfunction expansion. However, this method is not the only one; in Section 12.12 we showed how a solution can be constructed using the superposition integral. Here we expand upon this idea and illustrate how a solution, called a *Green's function*, to a differential equation forced by the Dirac delta function can be used in an integral representation of a solution when the forcing is arbitrary.

15.1 WHAT IS A GREEN'S FUNCTION?

The following examples taken from engineering show how Green's functions naturally appear during the solution of initial-value and boundary-value problems. We also show that the solution $u(x)$ can be expressed as an integral involving the Green's function and $f(x)$.

Circuit theory

In electrical engineering, one of the simplest electrical devices consists of a voltage source $v(t)$ connected to a resistor with resistance R and an inductor with inductance L. See Figure 15.1.1. Denoting the current by $i(t)$, the equation that governs this circuit is

$$L\frac{di}{dt} + Ri = v(t). \tag{15.1.1}$$

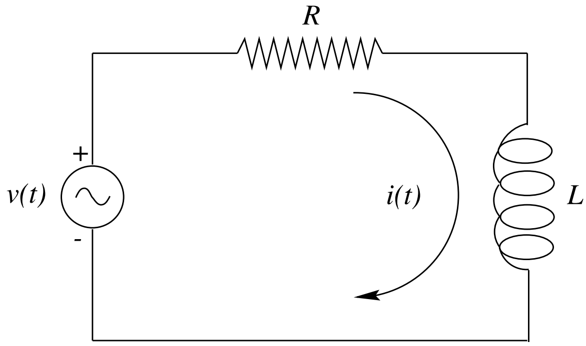


Figure 15.1.1: The RL electrical circuit driven by the voltage $v(t)$.

Consider now the following experiment: With the circuit initially dead, we allow the voltage to suddenly become $V_0/\Delta\tau$ during a very short duration $\Delta\tau$ starting at $t = \tau$. Then, at $t = \tau + \Delta\tau$, we again turn off the voltage supply. Mathematically, for $t > \tau + \Delta\tau$, the circuit's performance obeys the homogeneous differential equation

$$L\frac{di}{dt} + Ri = 0, \qquad t > \tau + \Delta\tau, \tag{15.1.2}$$

whose solution is

$$i(t) = I_0 e^{-Rt/L}, \qquad t > \tau + \Delta\tau, \tag{15.1.3}$$

where I_0 is a constant and L/R is the *time constant* of the circuit. Because the voltage $v(t)$ during $\tau < t < \tau + \Delta\tau$ is $V_0/\Delta\tau$, then

$$\int_\tau^{\tau+\Delta\tau} v(t)\,dt = V_0. \tag{15.1.4}$$

Therefore, over the interval $\tau < t < \tau + \Delta\tau$, Equation 15.1.1 can be integrated to yield

$$L\int_\tau^{\tau+\Delta\tau} di + R\int_\tau^{\tau+\Delta\tau} i(t)\,dt = \int_\tau^{\tau+\Delta\tau} v(t)\,dt, \tag{15.1.5}$$

or

$$L\left[i(\tau+\Delta\tau) - i(\tau)\right] + R\int_\tau^{\tau+\Delta\tau} i(t)\,dt = V_0. \tag{15.1.6}$$

If $i(t)$ remains continuous as $\Delta\tau$ becomes small, then

$$R\int_\tau^{\tau+\Delta\tau} i(t)\,dt \approx 0. \tag{15.1.7}$$

Finally, because

$$i(\tau) = 0, \tag{15.1.8}$$

and

$$i(\tau+\Delta\tau) = I_0 e^{-R(\tau+\Delta\tau)/L} \approx I_0 e^{-R\tau/L}, \tag{15.1.9}$$

for small $\Delta\tau$, Equation 15.1.6 reduces to

$$LI_0 e^{-R\tau/L} = V_0, \tag{15.1.10}$$

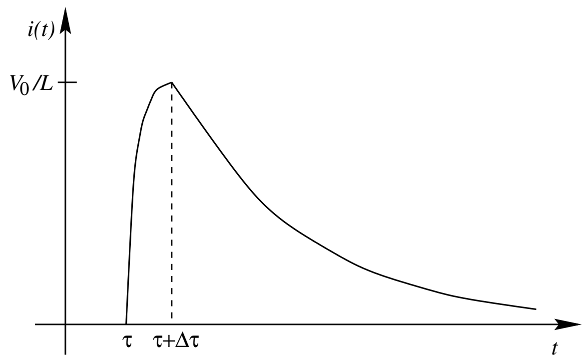


Figure 15.1.2: The current $i(t)$ within an RL circuit when the voltage $V_0/\Delta\tau$ is introduced between the times $\tau < t < \tau + \Delta\tau$.

or

$$I_0 = \frac{V_0}{L} e^{R\tau/L}. \tag{15.1.11}$$

Therefore, Equation 15.1.3 can be written as

$$i(t) = \begin{cases} 0, & t < \tau, \\ V_0 e^{-R(t-\tau)/L}/L, & \tau \le t, \end{cases} \tag{15.1.12}$$

after using Equation 15.1.11. Equation 15.1.12 is plotted in Figure 15.1.2.

Consider now a new experiment with the same circuit where we subject the circuit to N voltage impulses, each of duration $\Delta\tau$ and amplitude $V_i/\Delta\tau$ with $i = 0, 1, \ldots, N$, occurring at $t = \tau_i$. See Figure 15.1.3. The current response is then

$$i(t) = \begin{cases} 0, & t < \tau_0, \\ V_0 e^{-R(t-\tau_0)/L}/L, & \tau_0 < t < \tau_1, \\ V_0 e^{-R(t-\tau_0)/L}/L + V_1 e^{-R(t-\tau_1)/L}/L, & \tau_1 < t < \tau_2, \\ \vdots & \vdots \\ \displaystyle\sum_{i=0}^{N} V_i e^{-R(t-\tau_i)/L}/L, & \tau_N < t < \tau_{N+1}. \end{cases} \tag{15.1.13}$$

Finally, consider our circuit subjected to a continuous voltage source $v(t)$. Over each successive interval $d\tau$, the step change in voltage is $v(\tau)\,d\tau$. Consequently, from Equation 15.1.13 the response $i(t)$ is now given by the *superposition integral*

$$i(t) = \int_\tau^t \frac{v(\tau)}{L} e^{-R(t-\tau)/L}\,d\tau, \tag{15.1.14}$$

or

$$i(t) = \int_\tau^t v(\tau) g(t|\tau)\,d\tau, \tag{15.1.15}$$

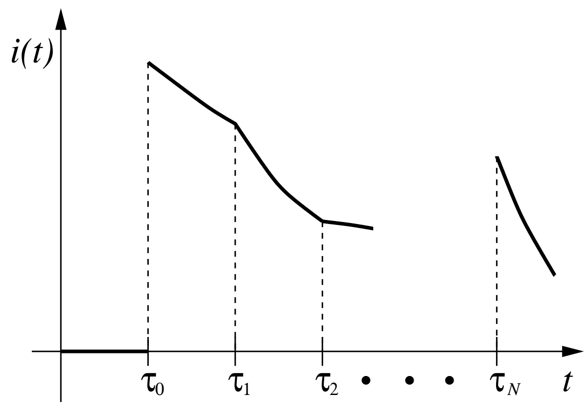


Figure 15.1.3: The current $i(t)$ within an RL circuit when the voltage is changed at $t = \tau_0$, $t = \tau_1$, and so forth.

where

$$g(t|\tau) = \frac{e^{-R(t-\tau)/L}}{L}, \qquad \tau < t. \tag{15.1.16}$$

Here we have assumed that $i(t) = v(t) = 0$ for $t < \tau$. In Equation 15.1.15, $g(t|\tau)$ is called the *Green's function*. As this equation shows, given the Green's function to Equation 15.1.1, the response $i(t)$ to any voltage source $v(t)$ can be obtained by convolving the voltage source with the Green's function.

We now show that we could have found the Green's function, Equation 15.1.16, by solving Equation 15.1.1 subject to an impulse- or delta-forcing function. Mathematically, this corresponds to solving the following initial-value problem:

$$L\frac{dg}{dt} + Rg = \delta(t-\tau), \qquad g(0|\tau) = 0. \tag{15.1.17}$$

Taking the Laplace transform of Equation 15.1.17, we find that

$$G(s|\tau) = \frac{e^{-s\tau}}{Ls + R}, \tag{15.1.18}$$

or

$$g(t|\tau) = \frac{e^{-R(t-\tau)/L}}{L} H(t-\tau), \tag{15.1.19}$$

where $H(\cdot)$ is the Heaviside step function. As our short derivation showed, the most direct route to finding a Green's function is solving the differential equation when its forcing equals the impulse or delta function. This is the technique that we will use throughout this chapter.

Statics

Consider a string of length L that is connected at both ends to supports and is subjected to a load (external force per unit length) of $f(x)$. We wish to find the displacement $u(x)$ of the string. If the load $f(x)$ acts downward (negative direction), the displacement $u(x)$ of the string is given by the differential equation

$$T\frac{d^2u}{dx^2} = f(x), \tag{15.1.20}$$

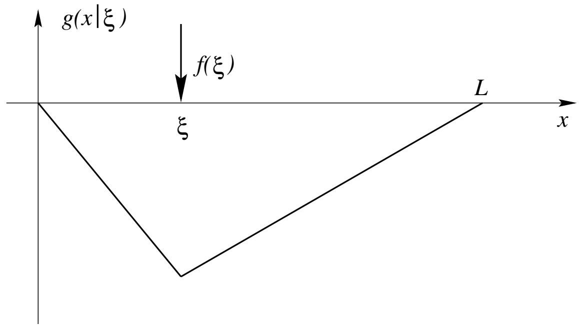


Figure 15.1.4: The response, commonly called a Green's function, of a string fixed at both ends to a point load at $x = \xi$.

where T denotes the uniform tensile force of the string. Because the string is stationary at both ends, the displacement $u(x)$ satisfies the boundary conditions

$$u(0) = u(L) = 0. \tag{15.1.21}$$

Instead of directly solving for the displacement $u(x)$ of the string subject to the load $f(x)$, let us find the displacement that results from a load $\delta(x-\xi)$ concentrated at the point $x = \xi$. See Figure 15.1.4. For this load, the differential equation, Equation 15.1.20, becomes

$$T\frac{d^2g}{dx^2} = \delta(x-\xi), \tag{15.1.22}$$

subject to the boundary conditions

$$g(0|\xi) = g(L|\xi) = 0. \tag{15.1.23}$$

In Equation 15.1.22, $g(x|\xi)$ denotes the displacement of the string when it is subjected to an impulse load at $x = \xi$. In line with our circuit theory example, it gives the *Green's function* for our statics problem. Once found, the displacement $u(x)$ of the string subject to any arbitrary load $f(x)$ can be found by convolving the load $f(x)$ with the Green's function $g(x|\xi)$ as we did earlier.

Let us now find this Green's function. At any point $x \neq \xi$, Equation 15.1.22 reduces to the homogeneous differential equation

$$\frac{d^2g}{dx^2} = 0, \tag{15.1.24}$$

which has the solution

$$g(x|\xi) = \begin{cases} ax+b, & 0 \le x < \xi, \\ cx+d, & \xi < x \le L. \end{cases} \tag{15.1.25}$$

Applying the boundary conditions, Equation 15.1.23, we find that

$$g(0|\xi) = a \cdot 0 + b = b = 0, \tag{15.1.26}$$

and

$$g(L|\xi) = cL + d = 0, \quad \text{or} \quad d = -cL. \tag{15.1.27}$$

Therefore, we can rewrite Equation 15.1.25 as

$$g(x|\xi) = \begin{cases} ax, & 0 \le x < \xi, \\ c(x-L), & \xi < x \le L, \end{cases} \tag{15.1.28}$$

where a and c are undetermined constants.

At $x = \xi$, the displacement $u(x)$ of the string must be continuous; otherwise, the string would be broken. Therefore, the Green's function given by Equation 15.1.28 must also be continuous there. Thus,

$$a\xi = c(\xi - L), \qquad \text{or} \qquad c = \frac{a\xi}{\xi - L}. \tag{15.1.29}$$

From Equation 15.1.22 the second derivative of $g(x|\xi)$ must equal the impulse function. Therefore, the first derivative of $g(x|\xi)$, obtained by integrating this equation, must be discontinuous by the amount $1/T$ or

$$\lim_{\epsilon\to 0}\left[\frac{dg(\xi+\epsilon|\xi)}{dx} - \frac{dg(\xi-\epsilon|\xi)}{dx}\right] = \frac{1}{T}, \tag{15.1.30}$$

in which case

$$\frac{dg(\xi^+|\xi)}{dx} - \frac{dg(\xi^-|\xi)}{dx} = \frac{1}{T}, \tag{15.1.31}$$

where ξ^+ and ξ^- denote points lying just above or below ξ, respectively. Using Equation 15.1.28, we find that

$$\frac{dg(\xi^-|\xi)}{dx} = a, \tag{15.1.32}$$

and

$$\frac{dg(\xi^+|\xi)}{dx} = c = \frac{a\xi}{\xi - L}. \tag{15.1.33}$$

Thus, Equation 15.1.31 leads to

$$\frac{a\xi}{\xi - L} - a = \frac{1}{T}, \quad \text{or} \quad \frac{aL}{\xi - L} = \frac{1}{T}, \quad \text{or} \quad a = \frac{\xi - L}{LT}, \tag{15.1.34}$$

and the Green's function is

$$g(x|\xi) = \frac{1}{TL}(x_> - L)x_<, \tag{15.1.35}$$

where $x_< = \min(x,\xi)$ and $x_> = \max(x,\xi)$. To find the displacement $u(x)$ subject to the load $f(x)$, we proceed as we did in the previous example. The result of this analysis is

$$u(x) = \int_0^L f(\xi)g(x|\xi)\,d\xi = \frac{x-L}{TL}\int_0^x f(\xi)\,\xi\,d\xi + \frac{x}{TL}\int_x^L f(\xi)\,(\xi - L)\,d\xi, \tag{15.1.36}$$

since $\xi < x$ in the first integral and $x < \xi$ in the second integral of Equation 15.1.36.

Integral Equations

Consider the Sturm-Liouville problem

$$y'' + \lambda y = 0, \qquad y(0) = y(L) = 0. \tag{15.1.37}$$

From Section 6.1, nontrivial solutions exist only if

$$\lambda_n = \frac{n^2\pi^2}{L^2}, \qquad y_n(x) = \sin\left(\frac{n\pi x}{L}\right), \tag{15.1.38}$$

where $n = 1, 2, 3, \ldots$.

Consider now a new boundary-value problem:

$$\frac{d^2y}{dx^2} = -f(x), \qquad y(0) = y(L) = 0. \tag{15.1.39}$$

In the next section (Equation 15.2.76), we will show that we can write its solution by

$$y(x) = \int_0^L f(\xi)g(x|\xi)\,d\xi, \tag{15.1.40}$$

where the Green's function $g(x|\xi)$ is given by

$$\frac{d^2g}{dx^2} = -\delta(x-\xi), \qquad g(0|\xi) = g(L|\xi) = 0, \tag{15.1.41}$$

or

$$g(x|\xi) = (L - x_>)x_</L, \tag{15.1.42}$$

where $x_> = \max(x, \xi)$ and $x_< = \min(x, \xi)$.

We can now use Equation 15.1.39 to rewrite Equation 15.1.37 as

$$\lambda y(\xi) = f(\xi). \tag{15.1.43}$$

Multiplying Equation 15.1.43 by $g(x|\xi)$ and integrating from 0 to L, we find that

$$\int_0^L f(\xi)g(x|\xi)\,d\xi = \lambda\int_0^L y(\xi)g(x|\xi)\,d\xi, \tag{15.1.44}$$

or

$$y(x) - \lambda\int_0^L y(\xi)g(x|\xi)\,d\xi = 0. \tag{15.1.45}$$

Because of the equivalence of Equation 15.1.37 and Equation 15.1.45, the solutions to the *integral equation* Equation 15.1.45 are $\lambda_n = n^2\pi^2/L^2$ with $y_n(x) = \sin(n\pi x/L)$. Direct substitution verifies this result. Thus, we can use Green's functions to construct integral equations that have known solutions. Indeed, it was the use of Green's functions to solve Fredholm integral equations that drew the attention of mathematicians at the turn of the twentieth century.[1]

[1] See Section 36 in Kneser, A., 1911: *Integralgleichungen und ihre Anwendungen in der mathematischen Physik*. Braunschweig, 293 pp.

15.2 ORDINARY DIFFERENTIAL EQUATIONS

Second-order differential equations are ubiquitous in engineering. In electrical engineering, many electrical circuits are governed by second-order, linear ordinary differential equations. In mechanical engineering they arise during the application of Newton's second law.

One of the drawbacks of solving ordinary differential equations with a forcing term is its lack of generality. Each new forcing function requires a repetition of the entire process. In this section we give some methods for finding the solution in a somewhat more general manner for stationary systems where the forcing, not any initially stored energy (i.e., nonzero initial conditions), produces the total output. Unfortunately, the solution must be written as an integral.

In Example 12.8.3 we solved the linear differential equation

$$y'' + 2y' + y = f(t), \tag{15.2.1}$$

subject to the initial conditions $y(0) = y'(0) = 0$. At that time we wrote the Laplace transform of $y(t)$, $Y(s)$, as the product of two Laplace transforms:

$$Y(s) = \frac{1}{(s+1)^2} F(s). \tag{15.2.2}$$

One drawback in using Equation 15.2.2 is its dependence upon an unspecified Laplace transform $F(s)$. Is there a way to eliminate this dependence and yet retain the essence of the solution?

One way of obtaining a quantity that is independent of the forcing is to consider the ratio:

$$\frac{Y(s)}{F(s)} = G(s) = \frac{1}{(s+1)^2}. \tag{15.2.3}$$

This ratio is called the *transfer function* because we can transfer the input $F(s)$ into the output $Y(s)$ by multiplying $F(s)$ by $G(s)$. It depends only upon the properties of the system.

Let us now consider a related problem to Equation 15.2.1, namely

$$g'' + 2g' + g = \delta(t), \qquad t > 0, \tag{15.2.4}$$

with $g(0) = g'(0) = 0$. Because the forcing equals the Dirac delta function, $g(t)$ is called the *impulse response* or *Green's function.*[2] Computing $G(s)$,

$$G(s) = \frac{1}{(s+1)^2}. \tag{15.2.5}$$

From Equation 15.2.3 we see that $G(s)$ is also the transfer function. Thus, an alternative method for computing the transfer function is to subject the system to impulse forcing and the Laplace transform of the response is the transfer function.

From Equation 15.2.3,

$$Y(s) = G(s)F(s), \tag{15.2.6}$$

[2] For the origin of the Green's function, see Farina, J. E. G., 1976: The work and significance of George Green, the miller mathematician, 1793–1841. *Bull. Inst. Math. Appl.*, **12**, 98–105.

or

$$y(t) = g(t) * f(t). \tag{15.2.7}$$

That is, the convolution of the impulse response with the particular forcing gives the response of the system. Thus, we may describe a stationary system in one of two ways: (1) in the transform domain we have the transfer function, and (2) in the time domain there is the impulse response.

Despite the fundamental importance of the impulse response or Green's function for a given linear system, it is often quite difficult to determine, especially experimentally, and a more convenient practice is to deal with the response to the unit step $H(t)$. This response is called the *indicial admittance* or *step response*, which we shall denote by $a(t)$. Because $\mathcal{L}[H(t)] = 1/s$, we can determine the transfer function from the indicial admittance because $\mathcal{L}[a(t)] = G(s)\mathcal{L}[H(t)]$ or $sA(s) = G(s)$. Furthermore, because

$$\mathcal{L}[g(t)] = G(s) = \frac{\mathcal{L}[a(t)]}{\mathcal{L}[H(t)]}, \tag{15.2.8}$$

then

$$g(t) = \frac{da(t)}{dt} \tag{15.2.9}$$

from Equation 12.1.18.

• Example 15.2.1

Let us find the transfer function, impulse response, and step response for the system

$$y'' - 3y' + 2y = f(t), \tag{15.2.10}$$

with $y(0) = y'(0) = 0$. To find the impulse response, we solve

$$g'' - 3g' + 2g = \delta(t - \tau), \tag{15.2.11}$$

with $g(0) = g'(0) = 0$. We have generalized the problem to an arbitrary forcing at $t = \tau$ and now denote the Green's function by $g(t|\tau)$. We have done this so that our discussion will be consistent with the other sections in the chapter.

Taking the Laplace transform of Equation 15.2.11, we find that

$$G(s|\tau) = \frac{e^{-s\tau}}{s^2 - 3s + 2}, \tag{15.2.12}$$

which is the transfer function for this system when $\tau = 0$. The impulse response or Green's function equals the inverse of $G(s|\tau)$ or

$$g(t|\tau) = \left[e^{2(t-\tau)} - e^{t-\tau}\right] H(t - \tau). \tag{15.2.13}$$

To find the step response, we solve

$$a'' - 3a' + 2a = H(t), \tag{15.2.14}$$

with $a(0) = a'(0) = 0$. Taking the Laplace transform of Equation 15.2.14,

$$A(s) = \frac{1}{s(s-1)(s-2)}, \tag{15.2.15}$$

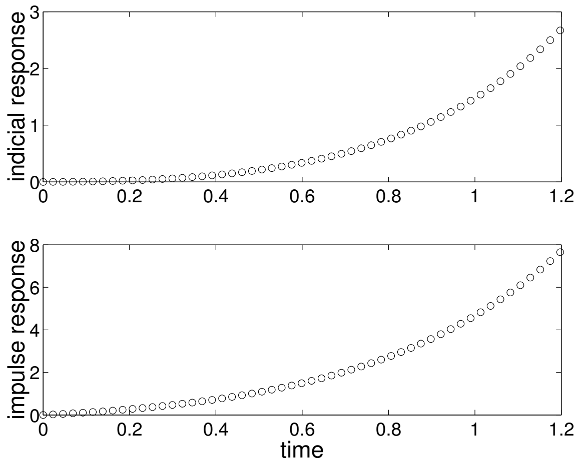


Figure 15.2.1: The impulse and step responses corresponding to the transfer function, Equation 15.2.12, with $\tau = 0$.

and the indicial admittance is given by the inverse of Equation 15.2.15, or

$$a(t) = \tfrac{1}{2} + \tfrac{1}{2}e^{2t} - e^t. \tag{15.2.16}$$

Note that $a'(t) = g(t|0)$. □

• Example 15.2.2

MATLAB's control toolbox contains several routines for the numerical computation of impulse and step responses if the transfer function can be written as the ratio of two polynomials. To illustrate this capacity, let us redo the previous example where the transfer function is given by Equation 15.2.12 with $\tau = 0$. The transfer function is introduced by loading in the polynomial in the numerator `num` and in the denominator `den` followed by calling `tf`. The MATLAB script

```
clear
% load in coefficients of the numerator and denominator
%     of the transfer function
num = [0 0 1]; den = [1 -3 2];
% create the transfer function
sys = tf(num,den);
% find the step response, a
[a,t] = step(sys);
% plot the indicial admittance
subplot(2,1,1), plot(t, a, 'o')
ylabel('indicial response','Fontsize',20)
% find the impulse response, g
[g,t] = impulse(sys);
% plot the impulse response
subplot(2,1,2), plot(t, g, 'o')
ylabel('impulse response','Fontsize',20)
xlabel('time','Fontsize',20)
```

shows how the impulse and step responses are found. Both of them are shown in Figure 15.2.1. □

• **Example 15.2.3**

There is an old joke about a man who took his car into a garage because of a terrible knocking sound. Upon his arrival the mechanic took one look at it and gave it a hefty kick.[3] Then, without a moment's hesitation he opened the hood, bent over, and tightened up a loose bolt. Turning to the owner, he said, "Your car is fine. That'll be $50." The owner felt that the charge was somewhat excessive, and demanded an itemized account. The mechanic said, "The kicking of the car and tightening one bolt, cost you a buck. The remaining $49 comes from knowing where to kick the car and finding the loose bolt."

Although the moral of the story may be about expertise as a marketable commodity, it also illustrates the concept of transfer function.[4] Let us model the car as a linear system where the equation

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_1 \frac{dy}{dt} + a_0 y = f(t) \tag{15.2.17}$$

governs the response $y(t)$ to a forcing $f(t)$. Assuming that the car has been sitting still, the initial conditions are zero and the Laplace transform of Equation 15.2.17 is

$$K(s)Y(s) = F(s), \tag{15.2.18}$$

where

$$K(s) = a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0. \tag{15.2.19}$$

Hence,

$$Y(s) = \frac{F(s)}{K(s)} = G(s)F(s), \tag{15.2.20}$$

where the transfer function $G(s)$ clearly depends only on the internal workings of the car. So if we know the transfer function, we understand how the car vibrates because

$$y(t) = \int_0^t g(t-x) f(x)\, dx. \tag{15.2.21}$$

But what does this have to do with our mechanic? He realized that a short sharp kick mimics an impulse forcing with $f(t) = \delta(t)$ and $y(t) = g(t)$. Therefore, by observing the response of the car to his kick, he diagnosed the loose bolt and fixed the car. □

In the previous examples, we used Laplace transforms to solve for the Green's functions. However, there is a rich tradition of using Fourier transforms rather than Laplace transforms. In these particular cases, the Fourier transform of the Green's function is called *frequency response* or *steady-state transfer function* of our system when $\tau = 0$. Consider the following examples.

---

[3] This is obviously a very old joke.

[4] Originally suggested by Stern, M. D., 1987: Why the mechanic kicked the car - A teaching aid for transfer functions. *Math. Gaz.*, **71**, 62–64.

• Example 15.2.4: Spectrum of a damped harmonic oscillator

In mechanics the damped oscillations of a mass m attached to a spring with a spring constant k and damped with a velocity-dependent resistance are governed by the equation

$$my'' + cy' + ky = f(t), \tag{15.2.22}$$

where $y(t)$ denotes the displacement of the oscillator from its equilibrium position, c denotes the damping coefficient, and $f(t)$ denotes the forcing.

Assuming that both $f(t)$ and $y(t)$ have Fourier transforms, let us analyze this system by finding its frequency response. We begin by solving for the Green's function $g(t|\tau)$, which is given by

$$mg'' + cg' + kg = \delta(t - \tau), \tag{15.2.23}$$

because the Green's function is the response of a system to a delta function forcing. Taking the Fourier transform of both sides of Equation 15.2.23, the frequency response is

$$G(\omega|\tau) = \frac{e^{-i\omega\tau}}{k + ic\omega - m\omega^2} = \frac{e^{-i\omega\tau}/m}{\omega_0^2 + ic\omega/m - \omega^2}, \tag{15.2.24}$$

where $\omega_0^2 = k/m$ is the natural frequency of the system. The most useful quantity to plot is the frequency response or

$$|G(\omega|\tau)| = \frac{\omega_0^2}{k\sqrt{(\omega^2 - \omega_0^2)^2 + \omega^2\omega_0^2(c^2/km)}} \tag{15.2.25}$$

$$= \frac{1}{k\sqrt{[(\omega/\omega_0)^2 - 1]^2 + (c^2/km)(\omega/\omega_0)^2}}. \tag{15.2.26}$$

In Figure 15.2.2 we plotted the frequency response as a function of $c^2/(km)$. Note that as the damping becomes larger, the sharp peak at $\omega = \omega_0$ essentially vanishes. As $c^2/(km) \to 0$, we obtain a very finely tuned response curve. Let us now find the Green's function. From the definition of the inverse Fourier transform,

$$mg(t|\tau) = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega t}}{\omega^2 - ic\omega/m - \omega_0^2}\,d\omega = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega t}}{(\omega - \omega_1)(\omega - \omega_2)}\,d\omega, \tag{15.2.27}$$

where

$$\omega_{1,2} = \pm\sqrt{\omega_0^2 - \gamma^2} + \gamma i, \tag{15.2.28}$$

and $\gamma = c/(2m) > 0$. We can evaluate Equation 15.2.27 by residues. Clearly the poles always lie in the upper half of the ω-plane. Thus, if $t < \tau$ in Equation 15.2.27 we can close the line integration along the real axis with a semicircle of infinite radius in the lower half of the ω-plane by Jordan's lemma. Because the integrand is analytic within the closed contour, $g(t|\tau) = 0$ for $t < \tau$. This is simply the causality condition,[5] the impulse forcing

[5] The principle stating that an event cannot precede its cause.

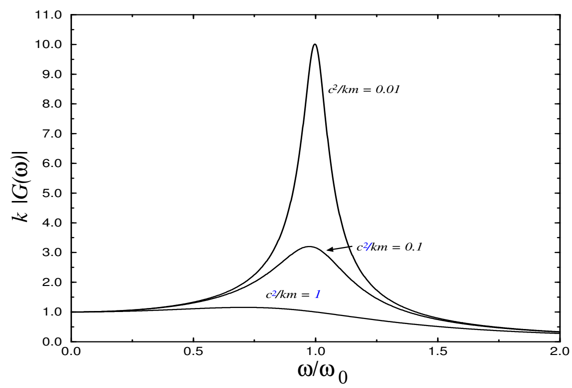


Figure 15.2.2: The variation of the frequency response for a damped harmonic oscillator as a function of driving frequency ω. See the text for the definition of the parameters.

being the cause of the excitation. Clearly, causality is closely connected with the analyticity of the frequency response in the lower half of the ω-plane.

If $t > \tau$, we close the line integration along the real axis with a semicircle of infinite radius in the upper half of the ω-plane and obtain

$$mg(t|\tau) = 2\pi i \left(-\frac{1}{2\pi}\right) \left\{ \text{Res}\left[\frac{e^{iz(t-\tau)}}{(z-\omega_1)(z-\omega_2)}; \omega_1\right] + \text{Res}\left[\frac{e^{iz(t-\tau)}}{(z-\omega_1)(z-\omega_2)}; \omega_2\right] \right\} \tag{15.2.29}$$

$$= \frac{-i}{\omega_1 - \omega_2} \left[e^{i\omega_1(t-\tau)} - e^{i\omega_2(t-\tau)}\right] = \frac{e^{-\gamma(t-\tau)} \sin\left[(t-\tau)\sqrt{\omega_0^2 - \gamma^2}\right]}{\sqrt{\omega_0^2 - \gamma^2}} H(t-\tau). \tag{15.2.30}$$

Let us now examine the damped harmonic oscillator by describing the migration of the poles $\omega_{1,2}$ in the complex ω-plane as γ increases from 0 to ∞. See Figure 15.2.3. For $\gamma \ll \omega_0$ (weak damping), the poles $\omega_{1,2}$ are very near to the real axis, above the points $\pm\omega_0$, respectively. This corresponds to the narrow resonance band discussed earlier and we have an underdamped harmonic oscillator. As γ increases from 0 to ω_0, the poles approach the positive imaginary axis, moving along a semicircle of radius ω_0 centered at the origin. They coalesce at the point $i\omega_0$ for $\gamma = \omega_0$, yielding repeated roots, and we have a critically damped oscillator. For $\gamma > \omega_0$, the poles move in opposite directions along the positive imaginary axis; one of them approaches the origin, while the other tends to $i\infty$ as $\gamma \to \infty$. The solution then has two purely decaying, overdamped solutions. During the early 1950s, a similar diagram was invented by Evans[6] where the movement of closed-loop poles is plotted for all values of a system parameter, usually the gain. This *root-locus method* is very popular in system control theory for two reasons. First, the investigator can easily determine the contribution of a particular closed-loop pole to the transient response. Second, he can determine the manner in which open-loop poles or zeros should be introduced or their location modified so that he will achieve a desired performance characteristic for his system. □

[6] Evans, W. R., 1948: Graphical analysis of control systems. *Trans. AIEE*, **67**, 547–551; Evans, W. R., 1954: *Control-System Dynamics*. McGraw-Hill, 282 pp.

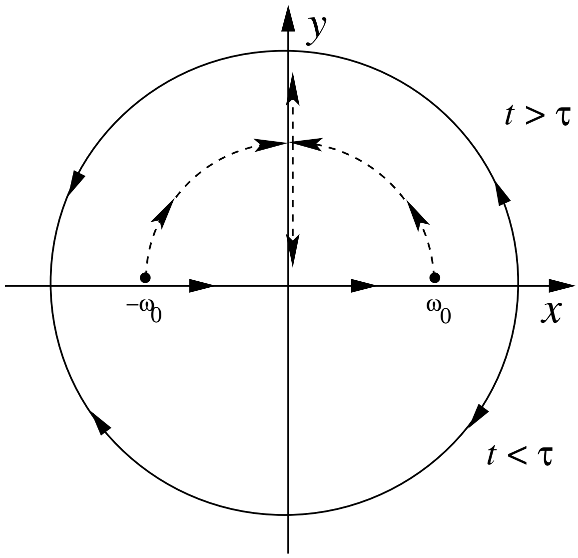


Figure 15.2.3: The migration of the poles of the frequency response of a damped harmonic oscillator as a function of γ.

• Example 15.2.5: Low-frequency filter

Consider the ordinary differential equation

$$Ry' + \frac{y}{C} = f(t), \tag{15.2.31}$$

where R and C are real, positive constants. If $y(t)$ denotes current, then Equation 15.2.31 would be the equation that gives the voltage across a capacitor in an RC circuit. Let us find the frequency response and Green's function for this system. We begin by writing Equation 15.2.31 as

$$Rg' + \frac{g}{C} = \delta(t-\tau), \tag{15.2.32}$$

where $g(t|\tau)$ denotes the Green's function. If the Fourier transform of $g(t|\tau)$ is $G(\omega|\tau)$, the frequency response $G(\omega|\tau)$ is given by

$$i\omega RG(\omega|\tau) + \frac{G(\omega|\tau)}{C} = e^{-i\omega\tau}, \tag{15.2.33}$$

or

$$G(\omega|\tau) = \frac{e^{-i\omega\tau}}{i\omega R + 1/C} = \frac{Ce^{-i\omega\tau}}{1+i\omega RC}, \tag{15.2.34}$$

and

$$|G(\omega|\tau)| = \frac{C}{\sqrt{1+\omega^2R^2C^2}} = \frac{C}{\sqrt{1+\omega^2/\omega_p^2}}, \tag{15.2.35}$$

where $\omega_p = 1/(RC)$ is an intrinsic constant of the system. In Figure 15.2.4 we plotted $|G(\omega|\tau)|$ as a function of ω. From this figure, we see that the response is largest for small ω and decreases as ω increases.

This is an example of a *low-frequency filter* because relatively more signal passes through at lower frequencies than at higher frequencies. To understand this, let us drive the system

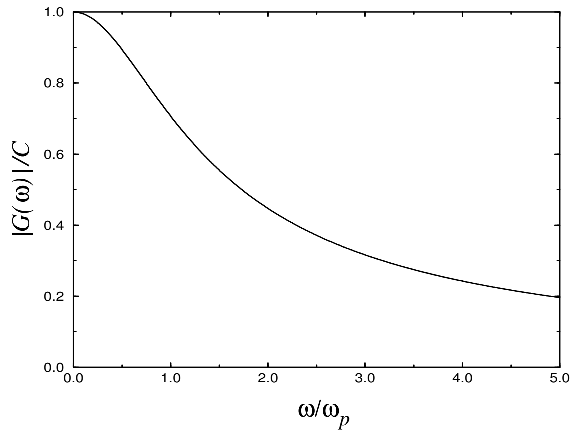


Figure 15.2.4: The variation of the frequency response, Equation 15.2.35, as a function of driving frequency ω. See the text for the definition of the parameters.

with a forcing function that has the Fourier transform $F(\omega)$. The response of the system will be $G(\omega,0)F(\omega)$. Thus, that portion of the forcing function's spectrum at the lower frequencies is relatively unaffected because $|G(\omega,0)|$ is near unity. However, at higher frequencies where $|G(\omega,0)|$ is smaller, the magnitude of the output is greatly reduced. □

• Example 15.2.6

During his study of tumor growth, Adam[7] found the particular solution to an ordinary differential equation which, in its simplest form, is

$$y'' - \alpha^2 y = \begin{cases} |x|/L - 1, & |x| < L, \\ 0, & |x| > L, \end{cases} \tag{15.2.36}$$

by the method of Green's functions. Let us retrace his steps and see how he did it.

The first step is finding the Green's function. We do this by solving

$$g'' - \alpha^2 g = \delta(x), \tag{15.2.37}$$

subject to the boundary conditions $\lim_{|x|\to\infty} g(x) \to 0$. Taking the Fourier transform of Equation 15.2.37, we obtain

$$G(\omega) = -\frac{1}{\omega^2 + \alpha^2}. \tag{15.2.38}$$

The function $G(\omega)$ is the *frequency response* for our problem. Straightforward inversion yields the Green's function

$$g(x) = -\frac{e^{-\alpha|x|}}{2\alpha}. \tag{15.2.39}$$

Therefore, by the convolution integral, $y(x) = g(x) * f(x)$,

$$y(x) = \int_{-L}^{L} g(x-\xi)\,(|\xi|/L - 1)\,d\xi = \frac{1}{2\alpha}\int_{-L}^{L} (1 - |\xi|/L)\,e^{-\alpha|x-\xi|}\,d\xi. \tag{15.2.40}$$

---

[7] Adam, J. A., 1986: A simplified mathematical model of tumor growth. *Math. Biosci.*, **81**, 229–244.

To evaluate Equation 15.2.40 we must consider four separate cases: $-\infty < x < -L$, $-L < x < 0$, $0 < x < L$, and $L < x < \infty$. Turning to the $-\infty < x < -L$ case first, we have

$$y(x) = \frac{1}{2\alpha}\int_{-L}^{L} (1 - |\xi|/L)\, e^{\alpha(x-\xi)}\, d\xi \tag{15.2.41}$$

$$= \frac{e^{\alpha x}}{2\alpha}\int_{-L}^{0} (1 + \xi/L)\, e^{-\alpha\xi}\, d\xi + \frac{e^{\alpha x}}{2\alpha}\int_{0}^{L} (1 - \xi/L)\, e^{-\alpha\xi}\, d\xi \tag{15.2.42}$$

$$= \frac{e^{\alpha x}}{2\alpha^3 L}\left(e^{\alpha L} + e^{-\alpha L} - 2\right). \tag{15.2.43}$$

Similarly, for $x > L$,

$$y(x) = \frac{1}{2\alpha}\int_{-L}^{L} (1 - |\xi|/L)\, e^{-\alpha(x-\xi)}\, d\xi \tag{15.2.44}$$

$$= \frac{e^{-\alpha x}}{2\alpha}\int_{-L}^{0} (1 + \xi/L)\, e^{\alpha\xi}\, d\xi + \frac{e^{-\alpha x}}{2\alpha}\int_{0}^{L} (1 - \xi/L)\, e^{\alpha\xi}\, d\xi \tag{15.2.45}$$

$$= \frac{e^{-\alpha x}}{2\alpha^3 L}\left(e^{\alpha L} + e^{-\alpha L} - 2\right). \tag{15.2.46}$$

On the other hand, for $-L < x < 0$, we find that

$$y(x) = \frac{1}{2\alpha}\int_{-L}^{x} (1 - |\xi|/L)\, e^{-\alpha(x-\xi)}\, d\xi + \frac{1}{2\alpha}\int_{x}^{L} (1 - |\xi|/L)\, e^{\alpha(x-\xi)}\, d\xi \tag{15.2.47}$$

$$= \frac{e^{-\alpha x}}{2\alpha}\int_{-L}^{x} (1 + \xi/L)\, e^{\alpha\xi}\, d\xi + \frac{e^{\alpha x}}{2\alpha}\int_{x}^{0} (1 + \xi/L)\, e^{-\alpha\xi}\, d\xi + \frac{e^{\alpha x}}{2\alpha}\int_{0}^{L} (1 - \xi/L)\, e^{-\alpha\xi}\, d\xi \tag{15.2.48}$$

$$= \frac{1}{\alpha^3 L}\left[e^{-\alpha L}\cosh(\alpha x) + \alpha(x + L) - e^{\alpha x}\right]. \tag{15.2.49}$$

Finally, for $0 < x < L$, we have that

$$y(x) = \frac{1}{2\alpha}\int_{-L}^{x} (1 - |\xi|/L)\, e^{-\alpha(x-\xi)}\, d\xi + \frac{1}{2\alpha}\int_{x}^{L} (1 - |\xi|/L)\, e^{\alpha(x-\xi)}\, d\xi \tag{15.2.50}$$

$$= \frac{e^{-\alpha x}}{2\alpha}\int_{-L}^{0} (1 + \xi/L)\, e^{\alpha\xi}\, d\xi + \frac{e^{-\alpha x}}{2\alpha}\int_{0}^{x} (1 - \xi/L)\, e^{\alpha\xi}\, d\xi + \frac{e^{\alpha x}}{2\alpha}\int_{x}^{L} (1 - \xi/L)\, e^{-\alpha\xi}\, d\xi \tag{15.2.51}$$

$$= \frac{1}{\alpha^3 L}\left[e^{-\alpha L}\cosh(\alpha x) + \alpha(L - x) - e^{-\alpha x}\right]. \tag{15.2.52}$$

These results can be collapsed down into

$$y(x) = \frac{1}{\alpha^3 L}\left[e^{-\alpha L}\cosh(\alpha x) + \alpha(L - |x|) - e^{-\alpha|x|}\right] \tag{15.2.53}$$

if $|x| < L$, and

$$y(x) = \frac{e^{-\alpha|x|}}{2\alpha^3 L}\left(e^{\alpha L} + e^{-\alpha L} - 2\right) \tag{15.2.54}$$

if $|x| > L$. □

Superposition integral

So far we showed how the response of any system can be expressed in terms of its Green's function and the arbitrary forcing. Can we also determine the response using the indicial admittance $a(t)$?

Consider first a system that is dormant until a certain time $t = \tau_1$. At that instant we subject the system to a forcing $H(t - \tau_1)$. Then the response will be zero if $t < \tau_1$ and will equal the indicial admittance $a(t - \tau_1)$ when $t > \tau_1$ because the indicial admittance is the response of a system to the step function. Here $t - \tau_1$ is the time measured from the instant of change.

Next, suppose that we now force the system with the value $f(0)$ when $t = 0$ and hold that value until $t = \tau_1$. We then abruptly change the forcing by an amount $f(\tau_1) - f(0)$ to the value $f(\tau_1)$ at the time τ_1 and hold it at that value until $t = \tau_2$. Then we again abruptly change the forcing by an amount $f(\tau_2) - f(\tau_1)$ at the time τ_2, and so forth (see Figure 15.2.5). From the *linearity* of the problem the response after the instant $t = \tau_n$ equals the sum

$$\begin{aligned} y(t) = f(0)a(t) + [f(\tau_1) - f(0)]a(t - \tau_1) + [f(\tau_2) - f(\tau_1)]a(t - \tau_2) \\ + \cdots + [f(\tau_n) - f(\tau_{n-1})]a(t - \tau_n). \end{aligned} \tag{15.2.55}$$

If we write $f(\tau_k) - f(\tau_{k-1}) = \Delta f_k$ and $\tau_k - \tau_{k-1} = \Delta\tau_k$, Equation 15.2.55 becomes

$$y(t) = f(0)a(t) + \sum_{k=1}^{n} a(t - \tau_k) \frac{\Delta f_k}{\Delta \tau_k}\, \Delta\tau_k. \tag{15.2.56}$$

Finally, proceeding to the limit as the number n of jumps becomes infinite, in such a manner that all jumps and intervals between successive jumps tend to zero, this sum has the limit

$$y(t) = f(0)a(t) + \int_0^t f'(\tau)a(t - \tau)\, d\tau. \tag{15.2.57}$$

Because the total response of the system equals the weighted sum (the weights being $a(t)$) of the forcing from the initial moment up to the time t, we refer to Equation 15.2.57 as the *superposition integral*, or *Duhamel's integral*,[8] named after the French mathematical physicist Jean-Marie-Constant Duhamel (1797–1872) who first derived it in conjunction with heat conduction.

We can also express Equation 15.2.57 in several different forms. Integration by parts yields

$$y(t) = f(t)a(0) + \int_0^t f(\tau)a'(t - \tau)\, d\tau = \frac{d}{dt}\left[\int_0^t f(\tau)a(t - \tau)\, d\tau\right]. \tag{15.2.58}$$

[8] Duhamel, J.-M.-C., 1833: Mémoire sur la méthode générale relative au mouvement de la chaleur dans les corps solides plongés dans des milieux dont la température varie avec le temps. *J. École Polytech.*, **22**, 20–77.

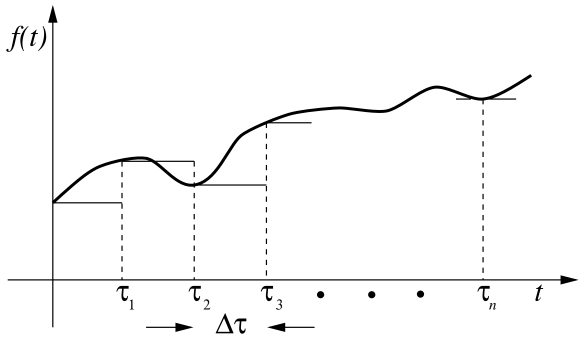


Figure 15.2.5: Diagram used in the derivation of Duhamel's integral.

• Example 15.2.7

Suppose that a system has the step response of $a(t) = A[1 - e^{-t/T}]$, where A and T are positive constants. Let us find the response if we force this system by $f(t) = kt$, where k is a constant.

From the superposition integral, Equation 15.2.57,

$$y(t) = 0 + \int_0^t kA[1 - e^{-(t-\tau)/T}]\,d\tau = kA[t - T(1 - e^{-t/T})]. \tag{15.2.59}$$

□

Boundary-value problem

One of the purposes of this book is the solution of a wide class of nonhomogeneous ordinary differential equations of the form

$$\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + s(x)y = -f(x), \qquad a \le x \le b, \tag{15.2.60}$$

with

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0, \qquad \beta_1 y(b) + \beta_2 y'(b) = 0. \tag{15.2.61}$$

For example, in Chapter 6 we examined the Sturm-Liouville-like equation

$$\frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + [q(x) + \lambda r(x)]y = -f(x), \qquad a \le x \le b, \tag{15.2.62}$$

where λ is a parameter. Here we wish to develop the Green's function for this class of boundary-value problems.

We begin by determining the Green's function for the equation

$$\frac{d}{dx}\left[p(x)\frac{dg}{dx}\right] + s(x)g = -\delta(x - \xi), \tag{15.2.63}$$

subject to yet undetermined boundary conditions. We know that such a function exists for the special case $p(x) = 1$ and $s(x) = 0$, and we now show that this is *almost always* true in the general case. Presently we construct Green's functions by requiring that they satisfy the following conditions:

- $g(x|\xi)$ satisfies the *homogeneous* equation $f(x) = 0$ *except* at $x = \xi$,
- $g(x|\xi)$ satisfies certain *homogeneous* conditions, and
- $g(x|\xi)$ is continuous at $x = \xi$.

These homogeneous boundary conditions for a finite interval (a, b) will be

$$\alpha_1 g(a|\xi) + \alpha_2 g'(a|\xi) = 0, \qquad \beta_1 g(b|\xi) + \beta_2 g'(b|\xi) = 0, \tag{15.2.64}$$

where g' denotes the x derivative of $g(x|\xi)$ and neither a nor b equals ξ. The coefficients α_1 and α_2 cannot both be zero; this also holds for β_1 and β_2. These conditions include the commonly encountered Dirichlet, Neumann, and Robin boundary conditions.

What about the value of $g'(x|\xi)$ at $x = \xi$? Because $g(x|\xi)$ is a continuous function of x, Equation 15.2.63 dictates that there must be a discontinuity in $g'(x|\xi)$ at $x = \xi$. We now show that this discontinuity consists of a jump in the value $g'(x|\xi)$ at $x = \xi$. To prove this, we begin by integrating Equation 15.2.63 from $\xi - \epsilon$ to $\xi + \epsilon$, which yields

$$p(x)\frac{dg(x|\xi)}{dx}\bigg|_{\xi-\epsilon}^{\xi+\epsilon} + \int_{\xi-\epsilon}^{\xi+\epsilon} s(x)g(x|\xi)\,dx = -1. \tag{15.2.65}$$

Because $g(x|\xi)$ and $s(x)$ are both continuous at $x = \xi$,

$$\lim_{\epsilon\to 0}\int_{\xi-\epsilon}^{\xi+\epsilon} s(x)g(x|\xi)\,dx = 0. \tag{15.2.66}$$

Applying the limit $\epsilon \to 0$ to Equation 15.2.65, we have that

$$p(\xi)\left[\frac{dg(\xi^+|\xi)}{dx} - \frac{dg(\xi^-|\xi)}{dx}\right] = -1, \tag{15.2.67}$$

where ξ^+ and ξ^- denote points just above and below $x = \xi$, respectively. Consequently, our last requirement on $g(x|\xi)$ will be that

- dg/dx must have a jump discontinuity of magnitude $-1/p(\xi)$ at $x = \xi$.

Similar conditions hold for higher-order ordinary differential equations.[9]

Consider now the region $a \le x < \xi$. Let $y_1(x)$ be a nontrivial solution of the *homogeneous* differential equation satisfying the boundary condition at $x = a$; then $\alpha_1 y_1(a) + \alpha_2 y_1'(a) = 0$. Because $g(x|\xi)$ must satisfy the same boundary condition, $\alpha_1 g(a|\xi) + \alpha_2 g'(a|\xi) = 0$. Since the set α_1, α_2 is nontrivial, then the Wronskian of y_1 and g must vanish at $x = a$

[9] Ince, E. L., 1956: *Ordinary Differential Equations*. Dover Publications, Inc. See Section 11.1.

or $y_1(a)g'(a|\xi) - y_1'(a)g(a|\xi) = 0$. However, for $a \le x < \xi$, both $y_1(x)$ and $g(x|\xi)$ satisfy the same differential equation, the homogeneous one. Therefore, their Wronskian is zero at all points and $g(x|\xi) = c_1 y_1(x)$ for $a \le x < \xi$, where c_1 is an arbitrary constant. In a similar manner, if the nontrivial function $y_2(x)$ satisfies the homogeneous equation and the boundary conditions at $x = b$, then $g(x|\xi) = c_2 y_2(x)$ for $\xi < x \le b$. The continuity condition of g and the jump discontinuity of g' at $x = \xi$ imply

$$c_1 y_1(\xi) - c_2 y_2(\xi) = 0, \qquad c_1 y_1'(\xi) - c_2 y_2'(\xi) = 1/p(\xi). \tag{15.2.68}$$

We can solve Equation 15.2.68 for c_1 and c_2 provided the Wronskian of y_1 and y_2 does not vanish at $x = \xi$, or

$$y_1(\xi)y_2'(\xi) - y_2(\xi)y_1'(\xi) \neq 0. \tag{15.2.69}$$

In other words, $y_1(x)$ must *not* be a multiple of $y_2(x)$. Is this always true? The answer is "generally yes." If the homogeneous equation admits no nontrivial solutions satisfying both boundary conditions at the same time,[10] then $y_1(x)$ and $y_2(x)$ must be linearly independent. On the other hand, if the homogeneous equation possesses a single solution, say $y_0(x)$, which also satisfies $\alpha_1 y_0(a) + \alpha_2 y_0'(a) = 0$ and $\beta_1 y_0(b) + \beta_2 y_0'(b) = 0$, then $y_1(x)$ will be a multiple of $y_0(x)$ and so is $y_2(x)$. Then they are multiples of each other and their Wronskian vanishes. This would occur, for example, if the differential equation is a Sturm-Liouville equation, λ equals the eigenvalue, and $y_0(x)$ is the corresponding eigenfunction. No Green's function exists in this case.

• Example 15.2.8

Consider the problem of finding the Green's function for $g'' + k^2 g = -\delta(x-\xi)$, $0 < x < L$, subject to the boundary conditions $g(0|\xi) = g(L|\xi) = 0$ with $k \neq 0$. The corresponding homogeneous equation is $y'' + k^2 y = 0$. Consequently, $g(x|\xi) = c_1 y_1(x) = c_1 \sin(kx)$ for $0 \le x \le \xi$, while $g(x|\xi) = c_2 y_2(x) = c_2 \sin[k(L-x)]$ for $\xi \le x \le L$.

Let us compute the Wronskian. For our particular problem,

$$W(x) = y_1(x)y_2'(x) - y_1'(x)y_2(x) \tag{15.2.70}$$

$$= -k\sin(kx)\cos[k(L-x)] - k\cos(kx)\sin[k(L-x)] \tag{15.2.71}$$

$$= -k\sin[k(x+L-x)] = -k\sin(kL), \tag{15.2.72}$$

and $W(\xi) = -k\sin(kL)$. Therefore, the Green's function will exist as long as $kL \neq n\pi$. If $kL = n\pi$, $y_1(x)$ and $y_2(x)$ are linearly *dependent* with $y_0(x) = c_3 \sin(n\pi x/L)$, the solution to the regular Sturm-Liouville problem $y'' + \lambda y = 0$, and $y(0) = y(L) = 0$. □

Let us now proceed to find $g(x|\xi)$ when it does exist. The system, Equation 15.2.68, has the unique solution

$$c_1 = -\frac{y_2(\xi)}{p(\xi)W(\xi)}, \quad \text{and} \quad c_2 = -\frac{y_1(\xi)}{p(\xi)W(\xi)}, \tag{15.2.73}$$

where $W(\xi)$ is the Wronskian of $y_1(x)$ and $y_2(x)$ at $x = \xi$. Therefore,

$$g(x|\xi) = -\frac{y_1(x_<)y_2(x_>)}{p(\xi)W(\xi)}. \tag{15.2.74}$$

[10] In the theory of differential equations, this system would be called *incompatible*: one that admits no solution, save $y = 0$, which is also continuous for all x in the interval (a, b) and satisfies the homogeneous boundary conditions.

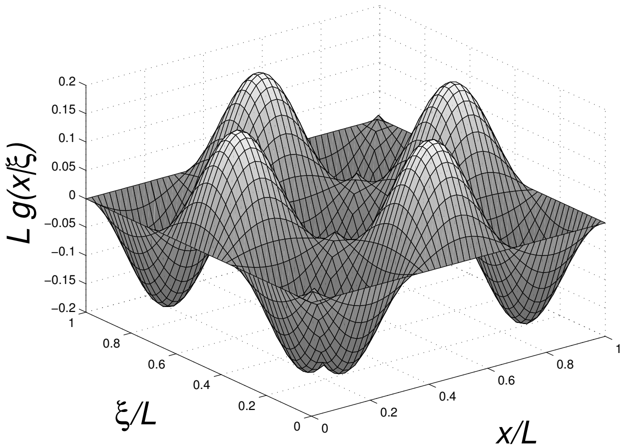


Figure 15.2.6: The Green's function, Equation 15.2.75, divided by L, as functions of x and ξ when $kL = 10$.

Clearly $g(x|\xi)$ is symmetric in x and ξ. It is also unique. The proof of the uniqueness is as follows: We can always choose a different $y_1(x)$, but it will be a multiple of the "old" $y_1(x)$, and the Wronskian will be multiplied by the same factor, leaving $g(x|\xi)$ the same. This is also true if we modify $y_2(x)$ in a similar manner.

• Example 15.2.9

Let us find the Green's function for $g'' + k^2 g = -\delta(x - \xi)$, $0 < x < L$, subject to the boundary conditions $g(0|\xi) = g(L|\xi) = 0$. As we showed in the previous example, $y_1(x) = c_1 \sin(kx)$, $y_2(x) = c_2 \sin[k(L - x)]$, and $W(\xi) = -k \sin(kL)$. Substituting into Equation 15.2.74, we have that

$$g(x|\xi) = \frac{\sin(kx_<)\sin[k(L - x_>)]}{k\,\sin(kL)}, \tag{15.2.75}$$

where $x_< = \min(x, \xi)$ and $x_> = \max(x, \xi)$. Figure 15.2.6 illustrates Equation 15.2.75. □

So far, we showed that the Green's function for Equation 15.2.63 exists, is symmetric, and enjoys certain properties (see the material in the boxes after Equation 15.2.63 and Equation 15.2.67). But how does this help us solve Equation 15.2.63? We now prove that

$$y(x) = \int_a^b g(x|\xi) f(\xi)\, d\xi \tag{15.2.76}$$

is the solution to the nonhomogeneous differential equation, Equation 15.2.63, and the homogeneous boundary conditions, Equation 15.2.64.

We begin by noting that in Equation 15.2.76 x is a parameter while ξ is the dummy variable. As we perform the integration, we must switch from the form for $g(x|\xi)$ for $\xi \le x$ to the second form for $\xi \ge x$ when ξ equals x; thus,

$$y(x) = \int_a^x g(x|\xi) f(\xi)\, d\xi + \int_x^b g(x|\xi) f(\xi)\, d\xi. \tag{15.2.77}$$

Differentiation yields

$$\frac{d}{dx}\int_a^x g(x|\xi)f(\xi)\,d\xi = \int_a^x \frac{dg(x|\xi)}{dx} f(\xi)\,d\xi + g(x|x^-)f(x), \tag{15.2.78}$$

and

$$\frac{d}{dx}\int_x^b g(x|\xi)f(\xi)\,d\xi = \int_x^b \frac{dg(x|\xi)}{dx} f(\xi)\,d\xi - g(x|x^+)f(x). \tag{15.2.79}$$

Because $g(x|\xi)$ is continuous everywhere, we have that $g(x|x^+) = g(x|x^-)$ so that

$$\frac{dy}{dx} = \int_a^x \frac{dg(x|\xi)}{dx} f(\xi)\,d\xi + \int_x^b \frac{dg(x|\xi)}{dx} f(\xi)\,d\xi. \tag{15.2.80}$$

Differentiating once more gives

$$\frac{d^2y}{dx^2} = \int_a^x \frac{d^2g(x|\xi)}{dx^2} f(\xi)\,d\xi + \frac{dg(x|x^-)}{dx} f(x) + \int_x^b \frac{d^2g(x|\xi)}{dx^2} f(\xi)\,d\xi - \frac{dg(x|x^+)}{dx} f(x). \tag{15.2.81}$$

The second and fourth terms on the right side of Equation 15.2.81 will not cancel in this case; on the contrary,

$$\frac{dg(x|x^-)}{dx} - \frac{dg(x|x^+)}{dx} = -\frac{1}{p(x)}. \tag{15.2.82}$$

To show this, we note that the term $dg(x|x^-)/dx$ denotes a differentiation of $g(x|\xi)$ with respect to x using the $x > \xi$ form and then letting $\xi \to x$. Thus,

$$\frac{dg(x|x^-)}{dx} = -\lim_{\substack{\xi\to x\\ \xi<x}} \frac{y_2'(x)y_1(\xi)}{p(\xi)W(\xi)} = -\frac{y_2'(x)y_1(x)}{p(x)W(x)}, \tag{15.2.83}$$

while for $dg(x|x^+)/dx$ we use the $x < \xi$ form or

$$\frac{dg(x|x^+)}{dx} = -\lim_{\substack{\xi\to x\\ \xi>x}} \frac{y_1'(x)y_2(\xi)}{p(\xi)W(\xi)} = -\frac{y_1'(x)y_2(x)}{p(x)W(x)}. \tag{15.2.84}$$

Upon introducing these results into the differential equation

$$p(x)\frac{d^2y}{dx^2} + p'(x)\frac{dy}{dx} + s(x)y = -f(x), \tag{15.2.85}$$

we have

$$\begin{aligned}\int_a^x &[p(x)g''(x|\xi) + p'(x)g'(x|\xi) + s(x)g(x|\xi)]f(\xi)\,d\xi \\ &+ \int_x^b [p(x)g''(x|\xi) + p'(x)g'(x|\xi) + s(x)g(x|\xi)]f(\xi)\,d\xi - p(x)\frac{f(x)}{p(x)} = -f(x).\end{aligned} \tag{15.2.86}$$

Because

$$p(x)g''(x|\xi) + p'(x)g'(x|\xi) + s(x)g(x|\xi) = 0, \tag{15.2.87}$$

except for $x = \xi$, Equation 15.2.86, and thus Equation 15.2.63, is satisfied. Although Equation 15.2.87 does not hold at the point $x = \xi$, the results are still valid because that one point does not affect the values of the integrals. As for the boundary conditions,

$$y(a) = \int_a^b g(a|\xi) f(\xi)\, d\xi, \qquad y'(a) = \int_a^b \frac{dg(a|\xi)}{dx} f(\xi)\, d\xi, \tag{15.2.88}$$

and $\alpha_1 y(a) + \alpha_2 y'(a) = 0$ from Equation 15.2.64. A similar proof holds for $x = b$.

Finally, let us consider the solution for the nonhomogeneous boundary conditions $\alpha_1 y(a) + \alpha_2 y'(a) = \alpha$, and $\beta_1 y(b) + \beta_2 y'(b) = \beta$. The solution in this case is

$$y(x) = \frac{\alpha y_2(x)}{\alpha_1 y_2(a) + \alpha_2 y_2'(a)} + \frac{\beta y_1(x)}{\beta_1 y_1(b) + \beta_2 y_1'(b)} + \int_a^b g(x|\xi) f(\xi)\, d\xi. \tag{15.2.89}$$

A quick check shows that Equation 15.2.89 satisfies the differential equation and both nonhomogeneous boundary conditions.

Eigenfunction expansion

We just showed how Green's functions can be used to solve the nonhomogeneous linear differential equation. The next question is how do you find the Green's function? Here we present the most common method: *series expansion*. This is not surprising given its success in solving the Sturm-Liouville problem.

Consider the nonhomogeneous problem

$$y'' = -f(x), \qquad \text{with} \qquad y(0) = y(L) = 0. \tag{15.2.90}$$

The Green's function $g(x|\xi)$ must therefore satisfy

$$g'' = -\delta(x - \xi), \quad \text{with} \quad g(0|\xi) = g(L|\xi) = 0. \tag{15.2.91}$$

Because $g(x|\xi)$ vanishes at the ends of the interval $(0, L)$, this suggests that it can be expanded in a series of suitably chosen orthogonal functions such as, for instance, the Fourier sine series

$$g(x|\xi) = \sum_{n=1}^{\infty} G_n(\xi) \sin\left(\frac{n\pi x}{L}\right), \tag{15.2.92}$$

where the expansion coefficients G_n are dependent on the parameter ξ. Although we chose the orthogonal set of functions $\sin(n\pi x/L)$, we could have used other orthogonal functions as long as they vanish at the endpoints.

Because

$$g''(x|\xi) = \sum_{n=1}^{\infty} \left(-\frac{n^2\pi^2}{L^2}\right) G_n(\xi) \sin\left(\frac{n\pi x}{L}\right), \tag{15.2.93}$$

and

$$\delta(x-\xi) = \sum_{n=1}^{\infty} A_n(\xi) \sin\left(\frac{n\pi x}{L}\right), \tag{15.2.94}$$

where

$$A_n(\xi) = \frac{2}{L}\int_0^L \delta(x-\xi)\sin\left(\frac{n\pi x}{L}\right)\,dx = \frac{2}{L}\sin\left(\frac{n\pi\xi}{L}\right), \tag{15.2.95}$$

we have that

$$-\sum_{n=1}^{\infty}\left(\frac{n^2\pi^2}{L^2}\right)G_n(\xi)\sin\left(\frac{n\pi x}{L}\right) = -\frac{2}{L}\sum_{n=1}^{\infty}\sin\left(\frac{n\pi\xi}{L}\right)\sin\left(\frac{n\pi x}{L}\right), \tag{15.2.96}$$

after substituting Equation 15.2.93 through Equation 15.2.95 into the differential equation, Equation 15.2.91. Since Equation 15.2.96 must hold for any arbitrary x,

$$\left(\frac{n^2\pi^2}{L^2}\right)G_n(\xi) = \frac{2}{L}\sin\left(\frac{n\pi\xi}{L}\right). \tag{15.2.97}$$

Thus, the Green's function is

$$g(x|\xi) = \frac{2L}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\sin\left(\frac{n\pi\xi}{L}\right)\sin\left(\frac{n\pi x}{L}\right). \tag{15.2.98}$$

How might we use Equation 15.2.98? We can use this series to construct the solution of the nonhomogeneous equation, Equation 15.2.90, via the formula

$$y(x) = \int_0^L g(x|\xi)\,f(\xi)\,d\xi. \tag{15.2.99}$$

This leads to

$$y(x) = \frac{2L}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\sin\left(\frac{n\pi x}{L}\right)\int_0^L f(\xi)\sin\left(\frac{n\pi\xi}{L}\right)\,d\xi, \tag{15.2.100}$$

or

$$y(x) = \frac{L^2}{\pi^2}\sum_{n=1}^{\infty}\frac{a_n}{n^2}\sin\left(\frac{n\pi x}{L}\right), \tag{15.2.101}$$

where a_n are the Fourier sine coefficients of $f(x)$.

• Example 15.2.10

Consider now the more complicated boundary-value problem

$$y'' + k^2 y = -f(x), \qquad \text{with} \qquad y(0) = y(L) = 0. \tag{15.2.102}$$

The Green's function $g(x|\xi)$ must now satisfy

$$g'' + k^2 g = -\delta(x-\xi), \quad \text{and} \quad g(0|\xi) = g(L|\xi) = 0. \tag{15.2.103}$$

Once again, we use the Fourier sine expansion

$$g(x|\xi) = \sum_{n=1}^{\infty} G_n(\xi) \sin\left(\frac{n\pi x}{L}\right). \quad \textbf{(15.2.104)}$$

Direct substitution of Equation 15.2.104 and Equation 15.2.94 into Equation 15.2.103 and grouping by corresponding harmonics yields

$$-\frac{n^2\pi^2}{L^2} G_n(\xi) + k^2 G_n(\xi) = -\frac{2}{L} \sin\left(\frac{n\pi\xi}{L}\right), \quad \textbf{(15.2.105)}$$

or

$$G_n(\xi) = \frac{2}{L} \frac{\sin(n\pi\xi/L)}{n^2\pi^2/L^2 - k^2}. \quad \textbf{(15.2.106)}$$

Thus, the Green's function is

$$g(x|\xi) = \frac{2}{L} \sum_{n=1}^{\infty} \frac{\sin(n\pi\xi/L)\sin(n\pi x/L)}{n^2\pi^2/L^2 - k^2}. \quad \textbf{(15.2.107)}$$

Examining Equation 15.2.107 more closely, we note that it enjoys the symmetry property that $g(x|\xi) = g(\xi|x)$. □

• **Example 15.2.11**

Let us find the series expansion for the Green's function for

$$xg'' + g' + \left(k^2 x - \frac{m^2}{x}\right) g = -\delta(x - \xi), \quad 0 < x < L, \quad \textbf{(15.2.108)}$$

where $m \geq 0$ and is an integer. The boundary conditions are

$$\lim_{x\to 0} |g(x|\xi)| < \infty, \quad \text{and} \quad g(L|\xi) = 0. \quad \textbf{(15.2.109)}$$

To find this series, consider the Fourier-Bessel series

$$g(x|\xi) = \sum_{n=1}^{\infty} G_n(\xi) J_m(k_{nm}x), \quad \textbf{(15.2.110)}$$

where k_{nm} is the nth root of $J_m(k_{nm}L) = 0$. This series enjoys the advantage that it satisfies the boundary conditions and we will not have to introduce any homogeneous solutions so that $g(x|\xi)$ satisfies the boundary conditions.

Substituting Equation 15.2.110 into Equation 15.2.108 after we divide by x and using the Fourier-Bessel expansion for the delta function, we have that

$$(k^2 - k_{nm}^2) G_n(\xi) = -\frac{2k_{nm}^2 J_m(k_{nm}\xi)}{L^2[J_{m+1}(k_{nm}L)]^2} = -\frac{2J_m(k_{nm}\xi)}{L^2[J_m'(k_{nm}L)]^2}, \quad \textbf{(15.2.111)}$$

so that

$$g(x|\xi) = \frac{2}{L^2} \sum_{n=1}^{\infty} \frac{J_m(k_{nm}\xi) J_m(k_{nm}x)}{(k_{nm}^2 - k^2)[J_m'(k_{nm}L)]^2}. \quad \textbf{(15.2.112)}$$

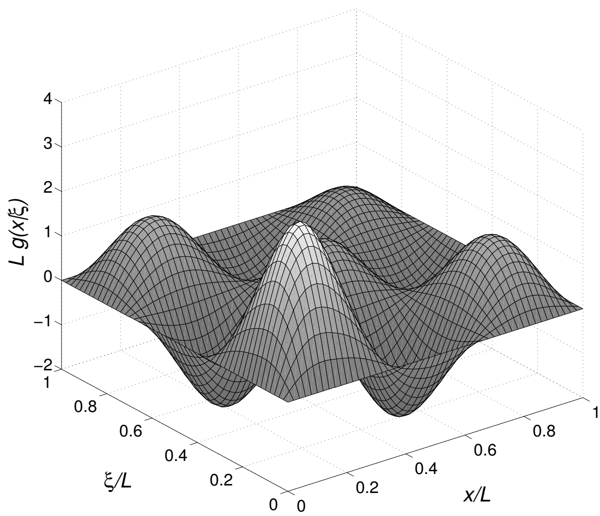


Figure 15.2.7: The Green's function, Equation 15.2.112, as functions of x/L and ξ/L when $kL = 10$ and $m = 1$.

Equation 15.2.112 is plotted in Figure 15.2.7. □

We summarize the expansion technique as follows: Suppose that we want to solve the differential equation

$$Ly(x) = -f(x), \qquad (15.2.113)$$

with some condition $By(x) = 0$ along the boundary, where L now denotes the Sturm-Liouville differential operator

$$L = \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + [q(x) + \lambda r(x)], \qquad (15.2.114)$$

and B is the boundary condition operator

$$B = \begin{cases} \alpha_1 + \alpha_2 \dfrac{d}{dx}, & \text{at } x = a, \\ \beta_1 + \beta_2 \dfrac{d}{dx}, & \text{at } x = b. \end{cases} \qquad (15.2.115)$$

We begin by seeking a Green's function $g(x|\xi)$, which satisfies

$$Lg = -\delta(x - \xi), \qquad Bg = 0. \qquad (15.2.116)$$

To find the Green's function, we utilize the set of eigenfunctions $\varphi_n(x)$ associated with the regular Sturm-Liouville problem

$$\frac{d}{dx}\left[p(x)\frac{d\varphi_n}{dx}\right] + [q(x) + \lambda_n r(x)]\varphi_n = 0, \qquad (15.2.117)$$

where $\varphi_n(x)$ satisfies the same boundary conditions as $y(x)$. If g exists and if the set $\{\varphi_n\}$ is complete, then $g(x|\xi)$ can be represented by the series

$$g(x|\xi) = \sum_{n=1}^{\infty} G_n(\xi)\varphi_n(x). \qquad \textbf{(15.2.118)}$$

Applying L to Equation 15.2.118,

$$Lg(x|\xi) = \sum_{n=1}^{\infty} G_n(\xi)L[\varphi_n(x)] = \sum_{n=1}^{\infty} G_n(\xi)(\lambda - \lambda_n)r(x)\varphi_n(x) = -\delta(x-\xi), \qquad \textbf{(15.2.119)}$$

if λ does not equal any of the eigenvalues λ_n. Multiplying both sides of Equation 15.2.119 by $\varphi_m(x)$ and integrating over x,

$$\sum_{n=1}^{\infty} G_n(\xi)(\lambda - \lambda_n) \int_a^b r(x)\varphi_n(x)\varphi_m(x)\,dx = -\varphi_m(\xi). \qquad \textbf{(15.2.120)}$$

If the eigenfunctions are *orthonormal*,

$$\int_a^b r(x)\varphi_n(x)\varphi_m(x)\,dx = \begin{cases} 1, & n = m, \\ 0, & n \neq m, \end{cases} \quad \text{and} \quad G_n(\xi) = \frac{\varphi_n(\xi)}{\lambda_n - \lambda}. \qquad \textbf{(15.2.121)}$$

This leads directly to the *bilinear formula*:

$$g(x|\xi) = \sum_{n=1}^{\infty} \frac{\varphi_n(\xi)\varphi_n(x)}{\lambda_n - \lambda}, \qquad \textbf{(15.2.122)}$$

which permits us to write the Green's function at once if the eigenvalues and eigenfunctions of L are known.

Problems

For the following initial-value problems, find the transfer function, impulse response, Green's function, and step response. Assume that all of the necessary initial conditions are zero and $\tau > 0$. If you have MATLAB's control toolbox, use MATLAB to check your work.

1. $g' + kg = \delta(t - \tau)$

2. $g'' - 2g' - 3g = \delta(t - \tau)$

3. $g'' + 4g' + 3g = \delta(t - \tau)$

4. $g'' - 2g' + 5g = \delta(t - \tau)$

5. $g'' - 3g' + 2g = \delta(t - \tau)$

6. $g'' + 4g' + 4g = \delta(t - \tau)$

7. $g'' - 9g = \delta(t - \tau)$

8. $g'' + g = \delta(t - \tau)$

9. $g'' - g' = \delta(t - \tau)$

Find the Green's function and the corresponding bilinear expansion using eigenfunctions from the regular Sturm-Liouville problem $\varphi_n'' + k_n^2\varphi_n = 0$ for

$$g'' = -\delta(x-\xi), \qquad 0 < x, \xi < L,$$

which satisfy the following boundary conditions:

10. $g(0|\xi) - \alpha g'(0|\xi) = 0, \alpha \neq 0, -L,$ $\qquad g(L|\xi) = 0,$

11. $g(0|\xi) - g'(0|\xi) = 0,$ $\qquad g(L|\xi) - g'(L|\xi) = 0,$

12. $g(0|\xi) - g'(0|\xi) = 0,$ $\qquad g(L|\xi) + g'(L|\xi) = 0.$

Find the Green's function[11] and the corresponding bilinear expansion using eigenfunctions from the regular Sturm-Liouville problem $\varphi_n'' + k_n^2\varphi_n = 0$ for

$$g'' - k^2 g = -\delta(x-\xi), \qquad 0 < x, \xi < L,$$

which satisfy the following boundary conditions:

13. $g(0|\xi) = 0,$ $\qquad g(L|\xi) = 0,$

14. $g'(0|\xi) = 0,$ $\qquad g'(L|\xi) = 0,$

15. $g(0|\xi) = 0,$ $\qquad g(L|\xi) + g'(L|\xi) = 0,$

16. $g(0|\xi) = 0,$ $\qquad g(L|\xi) - g'(L|\xi) = 0,$

17. $a\, g(0|\xi) + g'(0|\xi) = 0,$ $\qquad g'(L|\xi) = 0,$

18. $g(0|\xi) + g'(0|\xi) = 0,$ $\qquad g(L|\xi) - g'(L|\xi) = 0.$

15.3 JOINT TRANSFORM METHOD

In the previous section an important method for finding Green's function involved either Laplace or Fourier transforms. In the following sections we wish to find Green's functions for partial differential equations. Again transform methods play an important role. We will always use the Laplace transform to eliminate the temporal dependence. However, for the spatial dimension we will use either a Fourier series or Fourier transform. Our choice will be dictated by the domain: If it reaches to infinity, then we will employ Fourier transforms. On the other hand, a domain of finite length calls for an eigenfunction expansion. The following two examples illustrate our solution technique for domains of infinite and finite extent.

---

[11] Problem 18 was used by Chakrabarti, A., and T. Sahoo, 1996: Reflection of water waves by a nearly vertical porous wall. *J. Austral. Math. Soc., Ser. B*, **37**, 417–429.

• Example 15.3.1: One-dimensional Klein-Gordon equation

The Klein-Gordon equation is a form of the wave equation that arose in particle physics as the relativistic scalar wave equation describing particles with nonzero rest mass. In this example, we find its Green's function when there is only one spatial dimension:

$$\frac{\partial^2 g}{\partial x^2} - \frac{1}{c^2}\left(\frac{\partial^2 g}{\partial t^2} + a^2 g\right) = -\delta(x-\xi)\delta(t-\tau), \tag{15.3.1}$$

where $-\infty < x, \xi < \infty$, $0 < t, \tau$, c is a real, positive constant (the wave speed), and a is a real, nonnegative constant. The corresponding boundary conditions are

$$\lim_{|x|\to\infty} g(x,t|\xi,\tau) \to 0, \tag{15.3.2}$$

and the initial conditions are

$$g(x,0|\xi,\tau) = g_t(x,0|\xi,\tau) = 0. \tag{15.3.3}$$

We begin by taking the Laplace transform of Equation 15.3.1 and find that

$$\frac{d^2 G}{dx^2} - \left(\frac{s^2+a^2}{c^2}\right) G = -\delta(x-\xi)e^{-s\tau}. \tag{15.3.4}$$

Applying Fourier transforms to Equation 15.3.4, we obtain

$$G(x,s|\xi,\tau) = \frac{c^2}{2\pi}e^{-s\tau}\int_{-\infty}^{\infty} \frac{e^{ik(x-\xi)}}{s^2+a^2+k^2c^2}\,dk = \frac{c^2}{\pi}e^{-s\tau}\int_0^{\infty} \frac{\cos[k(x-\xi)]}{s^2+a^2+k^2c^2}\,dk. \tag{15.3.5}$$

Inverting the Laplace transform and employing the second shifting theorem,

$$g(x,t|\xi,\tau) = \frac{c^2}{\pi}H(t-\tau)\int_0^{\infty} \frac{\sin\left[(t-\tau)\sqrt{a^2+k^2c^2}\,\right]\cos[k(x-\xi)]}{\sqrt{a^2+k^2c^2}}\,dk. \tag{15.3.6}$$

Equation 15.3.6 represents a superposition of homogeneous solutions (normal modes) to Equation 15.3.1. An intriguing aspect of Equation 15.3.6 is that this solution occurs everywhere after $t > \tau$. If $|x-\xi| > c(t-\tau)$, these wave solutions destructively interfere so that we have zero there while they constructively interfere at those times and places where the physical waves are present.

Applying integral tables to Equation 15.3.6, the final result is

$$\boxed{g(x,t|\xi,\tau) = \frac{c}{2}J_0\left[a\sqrt{(t-\tau)^2-(x-\xi)^2/c^2}\,\right] H[c(t-\tau)-|x-\xi|].} \tag{15.3.7}$$

This Green's function is illustrated in Figure 15.3.1. Thus, the Green's function for the Klein-Gordon equation yields waves that propagate to the right and left from $x = 0$ with the wave front located at $x = \pm ct$. At a given point, after the passage of the wave front,

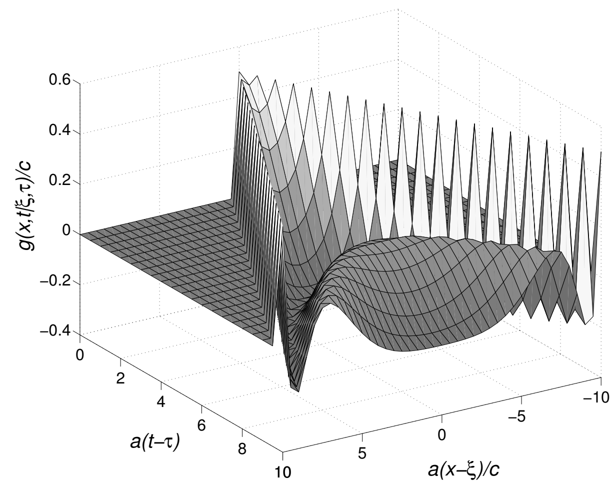


Figure 15.3.1: The free-space Green's function $g(x,t|\xi,\tau)/c$ for the one-dimensional Klein-Gordon equation at different distances $a(x-\xi)/c$ and times $a(t-\tau)$.

the solution vibrates with an ever-decreasing amplitude and at a frequency that approaches a - the so-called *cutoff frequency* - at $t \to \infty$.

Why is a called a cutoff frequency? From Equation 15.3.5, we see that, although the spectral representation includes all of the wavenumbers k running from $-\infty$ to ∞, the frequency $\omega = \sqrt{c^2k^2+a^2}$ is restricted to the range $\omega \ge a$ from Equation 15.3.6. Thus, a is the lowest possible frequency that a wave solution to the Klein-Gordon equation may have for a real value of k. □

• Example 15.3.2: One-dimensional wave equation on the interval $0 < x < L$

One of the classic problems of mathematical physics involves finding the displacement of a taut string between two supports when an external force is applied. The governing equation is

$$\frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} = f(x,t), \qquad 0<x<L, \quad 0<t, \tag{15.3.8}$$

where c is the constant phase speed.

In this example, we find the Green's function for this problem by considering the following problem:

$$\frac{\partial^2 g}{\partial t^2} - c^2\frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad 0<x,\xi<L, \quad 0<t,\tau, \tag{15.3.9}$$

with the boundary conditions

$$\alpha_1 g(0,t|\xi,\tau) + \beta_1 g_x(0,t|\xi,\tau) = 0, \quad 0<t, \tag{15.3.10}$$

and

$$\alpha_2 g(L,t|\xi,\tau) + \beta_2 g_x(L,t|\xi,\tau) = 0, \quad 0<t, \tag{15.3.11}$$

and the initial conditions

$$g(x,0|\xi,\tau) = g_t(x,0|\xi,\tau) = 0, \quad 0<x<L. \tag{15.3.12}$$

We start by taking the Laplace transform of Equation 15.3.9 and find that

$$\frac{d^2G}{dx^2} - \frac{s^2}{c^2}G = -\frac{\delta(x-\xi)}{c^2}e^{-s\tau}, \qquad 0 < x < L, \tag{15.3.13}$$

with

$$\alpha_1 G(0, s|\xi, \tau) + \beta_1 G'(0, s|\xi, \tau) = 0, \tag{15.3.14}$$

and

$$\alpha_2 G(L, s|\xi, \tau) + \beta_2 G'(L, s|\xi, \tau) = 0. \tag{15.3.15}$$

Problems similar to Equation 15.3.13 through Equation 15.3.15 were considered in the previous section. There, solutions were developed in terms of an eigenfunction expansion. Applying the same technique here,

$$G(x, s|\xi, \tau) = e^{-s\tau}\sum_{n=1}^{\infty}\frac{\varphi_n(\xi)\varphi_n(x)}{s^2 + c^2k_n^2}, \tag{15.3.16}$$

where $\varphi_n(x)$ is the nth *orthonormal* eigenfunction to the regular Sturm-Liouville problem

$$\varphi''(x) + k^2\varphi(x) = 0, \qquad 0 < x < L, \tag{15.3.17}$$

subject to the boundary conditions

$$\alpha_1\varphi(0) + \beta_1\varphi'(0) = 0, \tag{15.3.18}$$

and

$$\alpha_2\varphi(L) + \beta_2\varphi'(L) = 0. \tag{15.3.19}$$

Taking the inverse of Equation 15.3.16, we have that the Green's function is

$$\boxed{g(x, t|\xi, \tau) = \left\{\sum_{n=1}^{\infty}\varphi_n(\xi)\varphi_n(x)\frac{\sin[k_n c(t-\tau)]}{k_n c}\right\}H(t-\tau).} \tag{15.3.20}$$

Let us illustrate our results to find the Green's function for

$$\frac{\partial^2 g}{\partial t^2} - c^2\frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \tag{15.3.21}$$

with the boundary conditions

$$g(0, t|\xi, \tau) = g(L, t|\xi, \tau) = 0, \qquad 0 < t, \tag{15.3.22}$$

and the initial conditions

$$g(x, 0|\xi, \tau) = g_t(x, 0|\xi, \tau) = 0, \qquad 0 < x < L. \tag{15.3.23}$$

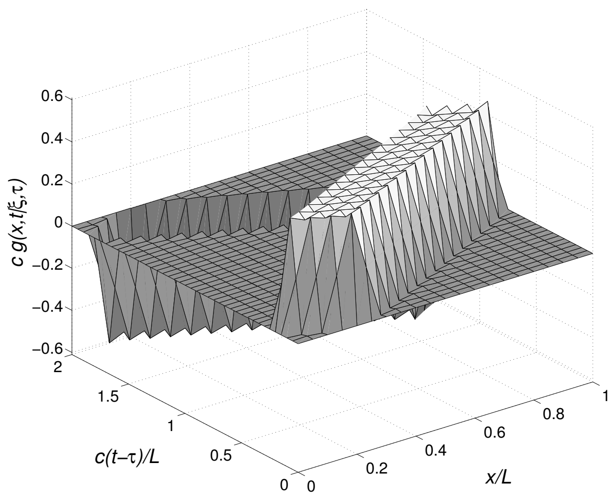


Figure 15.3.2: The Green's function $cg(x,t|\xi,\tau)$ given by Equation 15.3.26 for the one-dimensional wave equation over the interval $0 < x < L$ as a function of location x/L and time $c(t-\tau)/L$ with $\xi/L = 0.2$. The boundary conditions are $g(0,t|\xi,\tau) = g(L,t|\xi,\tau) = 0$.

For this example, the Sturm-Liouville problem is

$$\varphi''(x) + k^2\varphi(x) = 0, \qquad 0 < x < L, \tag{15.3.24}$$

with the boundary conditions $\varphi(0) = \varphi(L) = 0$. The nth *orthonormal* eigenfunction for this problem is

$$\varphi_n(x) = \sqrt{\frac{2}{L}}\,\sin\left(\frac{n\pi x}{L}\right). \tag{15.3.25}$$

Consequently, from Equation 15.3.20, the Green's function is

$$g(x,t|\xi,\tau) = \frac{2}{\pi c}\left\{\sum_{n=1}^{\infty}\frac{1}{n}\sin\left(\frac{n\pi\xi}{L}\right)\sin\left(\frac{n\pi x}{L}\right)\sin\left[\frac{n\pi c(t-\tau)}{L}\right]\right\}H(t-\tau). \tag{15.3.26}$$

See Figure 15.3.2.

15.4 WAVE EQUATION

In Section 15.2, we showed how Green's functions could be used to solve initial- and boundary-value problems involving ordinary differential equations. When we approach partial differential equations, similar considerations hold, although the complexity increases. In the next three sections, we work through the classic groupings of the wave, heat, and Helmholtz's equations in one spatial dimension. All of these results can be generalized to three dimensions.

Of these three groups, we start with the wave equation

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = -q(x,t), \tag{15.4.1}$$

where t denotes time, x is the position, c is the phase velocity of the wave, and $q(x,t)$ is the source density. In addition to Equation 15.4.1 it is necessary to state boundary and initial conditions to obtain a unique solution. The condition on the boundary can be either Dirichlet or Neumann or a linear combination of both (Robin condition). The conditions in time must be Cauchy - that is, we must specify the value of $u(x,t)$ and its time derivative at $t = t_0$ for each point of the region under consideration.

We begin by proving that we can express the solution to Equation 15.4.1 in terms of boundary conditions, initial conditions, and the Green's function, which is found by solving

$$\frac{\partial^2 g}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 g}{\partial t^2} = -\delta(x-\xi)\delta(t-\tau), \qquad \textbf{(15.4.2)}$$

where ξ denotes the position of a source that is excited at $t = \tau$. Equation 15.4.2 expresses the effect of an impulse as it propagates from $x = \xi$ as time increases from $t = \tau$. For $t < \tau$, causality requires that $g(x,t|\xi,\tau) = g_t(x,t|\xi,\tau) = 0$ if the impulse is the sole source of the disturbance. We also require that g satisfies the homogeneous form of the boundary condition satisfied by u.

Our derivation starts with the equations

$$\frac{\partial^2 u(\xi,\tau)}{\partial \xi^2} - \frac{1}{c^2}\frac{\partial^2 u(\xi,\tau)}{\partial \tau^2} = -q(\xi,\tau), \qquad \textbf{(15.4.3)}$$

and

$$\frac{\partial^2 g(x,t|\xi,\tau)}{\partial \xi^2} - \frac{1}{c^2}\frac{\partial^2 g(x,t|\xi,\tau)}{\partial \tau^2} = -\delta(x-\xi)\delta(t-\tau), \qquad \textbf{(15.4.4)}$$

where we obtain Equation 15.4.4 from a combination of Equation 15.4.2 plus reciprocity, namely $g(x,t|\xi,\tau) = g(\xi,-\tau|x,-t)$. Next we multiply Equation 15.4.3 by $g(x,t|\xi,\tau)$ and Equation 15.4.4 by $u(\xi,\tau)$ and subtract. Integrating over ξ from a to b, where a and b are the endpoints of the spatial domain, and over τ from 0 to t^+, where t^+ denotes a time slightly later than t so that we avoid ending the integration exactly at the peak of the delta function, we obtain

$$\begin{aligned}\int_0^{t^+}\int_a^b \Bigg\{ g(x,t|\xi,\tau)\frac{\partial^2 u(\xi,\tau)}{\partial \xi^2} &- u(\xi,\tau)\frac{\partial^2 g(x,t|\xi,\tau)}{\partial \xi^2} \\ &+ \frac{1}{c^2}\left[u(\xi,\tau)\frac{\partial^2 g(x,t|\xi,\tau)}{\partial \tau^2} - g(x,t|\xi,\tau)\frac{\partial^2 u(\xi,\tau)}{\partial \tau^2}\right]\Bigg\}\,d\xi\,d\tau \\ &= u(x,t) - \int_0^{t^+}\int_a^b q(\xi,\tau)\; g(x,t|\xi,\tau)\,d\xi\,d\tau. \qquad \textbf{(15.4.5)}\end{aligned}$$

Because

$$\begin{aligned} g(x,t|\xi,\tau)\frac{\partial^2 u(\xi,\tau)}{\partial \xi^2} &- u(\xi,\tau)\frac{\partial^2 g(x,t|\xi,\tau)}{\partial \xi^2} \\ &= \frac{\partial}{\partial \xi}\left[g(x,t|\xi,\tau)\frac{\partial u(\xi,\tau)}{\partial \xi}\right] - \frac{\partial}{\partial \xi}\left[u(\xi,\tau)\frac{\partial g(x,t|\xi,\tau)}{\partial \xi}\right], \qquad \textbf{(15.4.6)}\end{aligned}$$

and

$$\begin{aligned} g(x,t|\xi,\tau)\frac{\partial^2 u(\xi,\tau)}{\partial \tau^2} &- u(\xi,\tau)\frac{\partial^2 g(x,t|\xi,\tau)}{\partial \tau^2} \\ &= \frac{\partial}{\partial \tau}\left[g(x,t|\xi,\tau)\frac{\partial u(\xi,\tau)}{\partial \tau}\right] - \frac{\partial}{\partial \tau}\left[u(\xi,\tau)\frac{\partial g(x,t|\xi,\tau)}{\partial \tau}\right], \qquad \textbf{(15.4.7)}\end{aligned}$$

we find that

$$\int_0^{t^+}\left[g(x,t|\xi,\tau)\frac{\partial u(\xi,\tau)}{\partial \xi} - u(\xi,\tau)\frac{\partial g(x,t|\xi,\tau)}{\partial \xi}\right]_{\xi=a}^{\xi=b} d\tau$$
$$+\frac{1}{c^2}\int_a^b\left[u(\xi,\tau)\frac{\partial g(x,t|\xi,\tau)}{\partial \tau} - g(x,t|\xi,\tau)\frac{\partial u(\xi,\tau)}{\partial \tau}\right]_{\tau=0}^{\tau=t^+} d\xi$$
$$+\int_0^{t^+}\int_a^b q(\xi,\tau)\; g(x,t|\xi,\tau)\,d\xi\,d\tau = u(x,t). \tag{15.4.8}$$

The integrand in the first integral is specified by the boundary conditions. In the second integral, the integrand vanishes at $t = t^+$ from the initial conditions on $g(x,t|\xi,\tau)$. The limit at $t = 0$ is determined by the initial conditions. Hence,

$$u(x,t) = \int_0^{t^+}\int_a^b q(\xi,\tau) g(x,t|\xi,\tau)\,d\xi\,d\tau$$
$$+\int_0^{t^+}\left[g(x,t|\xi,\tau)\frac{\partial u(\xi,\tau)}{\partial \xi} - u(\xi,\tau)\frac{\partial g(x,t|\xi,\tau)}{\partial \xi}\right]_{\xi=a}^{\xi=b} d\tau$$
$$-\frac{1}{c^2}\int_a^b\left[u(\xi,0)\frac{\partial g(x,t|\xi,0)}{\partial \tau} - g(x,t|\xi,0)\frac{\partial u(\xi,0)}{\partial \tau}\right] d\xi. \tag{15.4.9}$$

Equation 15.4.9 gives the complete solution of the nonhomogeneous problem. The first two integrals on the right side of this equation represent the effect of the source and the boundary conditions, respectively. The last term involves the initial conditions; it can be interpreted as asking what sort of source is needed so that the function $u(x,t)$ starts in the desired manner.

• Example 15.4.1

Let us apply the Green's function technique to solve

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < 1, \quad 0 < t, \tag{15.4.10}$$

subject to the boundary conditions $u(0,t) = 0$ and $u(1,t) = t$, $0 < t$, and the initial conditions $u(x,0) = x$ and $u_t(x,0) = 0$, $0 < x < 1$.

Because there is no source term and $c = 1$, Equation 15.4.9 becomes

$$u(x,t) = \int_0^t [g(x,t|1,\tau)u_\xi(1,\tau) - u(1,\tau)g_\xi(x,t|1,\tau)]\; d\tau$$
$$-\int_0^t [g(x,t|0,\tau)u_\xi(0,\tau) - u(0,\tau)g_\xi(x,t|0,\tau)]\; d\tau$$
$$-\int_0^1 [u(\xi,0)g_\tau(x,t|\xi,0) - g(x,t|\xi,0)u_\xi(\xi,0)]\; d\xi. \tag{15.4.11}$$

Therefore we must first compute the Green's function for this problem. However, we have already done this in Example 15.3.2 and it is given by Equation 15.3.26 with $c = L = 1$.

Next, we note that $g(x,t|1,\tau) = g(x,t|0,\tau) = 0$ and $u(0,\tau) = u_\tau(\xi,0) = 0$. Consequently, Equation 15.4.11 reduces to only two nonvanishing integrals:

$$u(x,t) = -\int_0^t u(1,\tau) g_\xi(x,t|1,\tau)\,d\tau - \int_0^1 u(\xi,0) g_\tau(x,t|\xi,0)\,d\xi. \tag{15.4.12}$$

If we now substitute for $g(x,t|\xi,\tau)$ and reverse the order of integration and summation,

$$\int_0^t u(1,\tau) g_\xi(x,t|1,\tau)\,d\tau = 2\sum_{n=1}^\infty (-1)^n \sin(n\pi x) \int_0^t \tau \sin[n\pi(t-\tau)]\,d\tau \tag{15.4.13}$$

$$= 2t\sum_{n=1}^\infty (-1)^n \sin(n\pi x) \int_0^t \sin[n\pi(t-\tau)]\,d(t-\tau)$$

$$- 2\sum_{n=1}^\infty (-1)^n \sin(n\pi x) \int_0^t (t-\tau)\sin[n\pi(t-\tau)]\,d(t-\tau) \tag{15.4.14}$$

$$= -2t\sum_{n=1}^\infty (-1)^n \sin(n\pi x) \left.\frac{\cos[n\pi(t-\tau)]}{n\pi}\right|_0^t \tag{15.4.15}$$

$$- 2\sum_{n=1}^\infty (-1)^n \sin(n\pi x) \left.\left\{\frac{\sin[n\pi(t-\tau)]}{n^2\pi^2} - (t-\tau)\frac{\cos[n\pi(t-\tau)]}{n\pi}\right\}\right|_0^t$$

$$= -\frac{2t}{\pi}\sum_{n=1}^\infty \frac{(-1)^n}{n}\sin(n\pi x) + \frac{2}{\pi^2}\sum_{n=1}^\infty \frac{(-1)^n}{n^2}\sin(n\pi x)\sin(n\pi t), \tag{15.4.16}$$

and

$$\int_0^1 u(\xi,0) g_\tau(x,t|\xi,0)\,d\xi = -2\sum_{n=1}^\infty \sin(n\pi x)\cos(n\pi t) \int_0^1 \xi \sin(n\pi\xi)\,d\xi \tag{15.4.17}$$

$$= -2\sum_{n=1}^\infty \sin(n\pi x)\cos(n\pi t) \left.\left[\frac{\sin(n\pi\xi)}{n^2\pi^2} - \frac{\xi\,\cos(n\pi\xi)}{n\pi}\right]\right|_0^1 \tag{15.4.18}$$

$$= \frac{2}{\pi}\sum_{n=1}^\infty \frac{(-1)^n}{n}\sin(n\pi x)\cos(n\pi t). \tag{15.4.19}$$

Substituting Equation 15.4.16 and Equation 15.4.19 into Equation 15.4.12, we finally obtain

$$u(x,t) = -\frac{2t}{\pi}\sum_{n=1}^\infty \frac{(-1)^n}{n}\sin(n\pi x) - \frac{2}{\pi}\sum_{n=1}^\infty \frac{(-1)^n}{n}\sin(n\pi x)\cos(n\pi t)$$

$$+ \frac{2}{\pi^2}\sum_{n=1}^\infty \frac{(-1)^n}{n^2}\sin(n\pi x)\sin(n\pi t). \tag{15.4.20}$$

The first summation in Equation 15.4.20 is the Fourier sine expansion for $f(x) = x$ over the interval $0 < x < 1$. Indeed, a quick check shows that the particular solution $u_p(x,t) = xt$ satisfies the partial differential equation and boundary conditions. The remaining two summations are necessary so that $u(x,0) = x$ and $u_t(x,0) = 0$. □

To apply Equation 15.4.9 to other problems, we must now find the Green's function for a specific domain. In the following examples we illustrate how this is done using the joint transform method introduced in the previous section. Note that both examples given there were for the wave equation.

• Example 15.4.2: One-dimensional wave equation in an unlimited domain

The simplest possible example of Green's functions for the wave equation is the one-dimensional vibrating string problem.[12] In this problem the Green's function is given by the equation

$$\frac{\partial^2 g}{\partial t^2} - c^2 \frac{\partial^2 g}{\partial x^2} = c^2 \delta(x-\xi)\delta(t-\tau), \tag{15.4.21}$$

where $-\infty < x, \xi < \infty$, and $0 < t, \tau$. If the initial conditions equal zero, the Laplace transform of Equation 15.4.21 is

$$\frac{d^2 G}{dx^2} - \frac{s^2}{c^2} G = -\delta(x-\xi)e^{-s\tau}, \tag{15.4.22}$$

where $G(x, s|\xi, \tau)$ is the Laplace transform of $g(x, t|\xi, \tau)$. To solve Equation 15.4.22 we take its Fourier transform and obtain the algebraic equation

$$\overline{G}(k, s|\xi, \tau) = \frac{\exp(-ik\xi - s\tau)}{k^2 + s^2/c^2}. \tag{15.4.23}$$

Having found the joint Laplace-Fourier transform of $g(x, t|\xi, \tau)$, we must work our way back to the Green's function. From the definition of the Fourier transform, we have that

$$G(x, s|\xi, \tau) = \frac{e^{-s\tau}}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ik(x-\xi)}}{k^2 + s^2/c^2}\, dk. \tag{15.4.24}$$

To evaluate the Fourier-type integral, Equation 15.4.24, we apply the residue theorem. See Section 11.4. Performing the calculation,

$$G(x, s|\xi, \tau) = \frac{c\ \exp(-s\tau - s|x-\xi|/c)}{2s}. \tag{15.4.25}$$

Finally, taking the inverse Laplace transform of Equation 15.4.25,

$$g(x, t|\xi, \tau) = \frac{c}{2} H\left(t - \tau - |x-\xi|/c\right), \tag{15.4.26}$$

or

$$\boxed{g(x, t|\xi, \tau) = \frac{c}{2} H\left[c(t-\tau) + (x-\xi)\right] H\left[c(t-\tau) - (x-\xi)\right].} \tag{15.4.27}$$

We can use Equation 15.4.26 and the *method of images* to obtain the Green's function for

$$\frac{\partial^2 g}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 g}{\partial t^2} = \delta(x-\xi)\delta(t-\tau), \qquad 0 < x, t, \xi, \tau, \tag{15.4.28}$$

[12] See also Graff, K. F., 1991: *Wave Motion in Elastic Solids*. Dover Publications, Inc., Section 1.1.8.

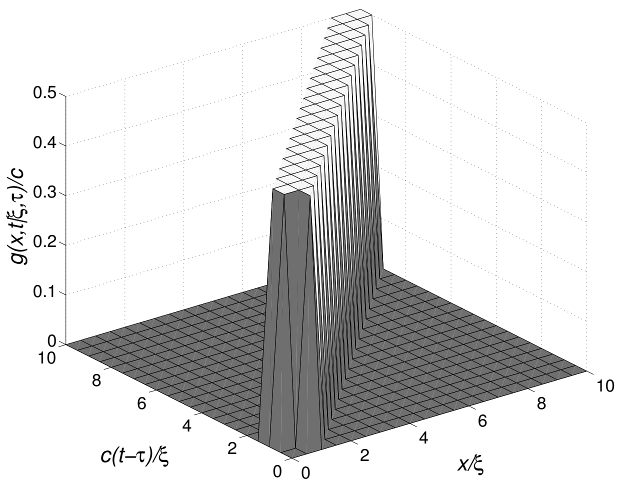


Figure 15.4.1: The Green's function $g(x,t|\xi,\tau)/c$ given by Equation 15.4.30 for the one-dimensional wave equation for $x > 0$ at different distances x/ξ and times $c(t-\tau)$ subject to the boundary condition $g(0,t|\xi,\tau) = 0$.

subject to the boundary condition $g(0,t|\xi,\tau) = 0$.

We begin by noting that the free-space Green's function,[13] Equation 15.4.26, is the particular solution to Equation 15.4.28. Therefore, we need only find a homogeneous solution $f(x,t|\xi,\tau)$ so that

$$g(x,t|\xi,\tau) = \frac{c}{2} H\left(t - \tau - |x-\xi|/c\right) + f(x,t|\xi,\tau) \qquad \textbf{(15.4.29)}$$

satisfies the boundary condition at $x = 0$.

To find $f(x,t|\xi,\tau)$, let us introduce a source at $x = -\xi$ at $t = \tau$. The corresponding free-space Green's function is $H(t-\tau-|x+\xi|/c)$. If, along the boundary $x = 0$ for any time t, this Green's function destructively interferes with the free-space Green's function associated with the source at $x = \xi$, then we have our solution. This will occur if our new source has a negative sign, resulting in the combined Green's function

$$g(x,t|\xi,\tau) = \frac{c}{2}\left[H(t-\tau-|x-\xi|/c) - H(t-\tau-|x+\xi|/c)\right]. \qquad \textbf{(15.4.30)}$$

See Figure 15.4.1. Because Equation 15.4.30 satisfies the boundary condition, we need no further sources.

In a similar manner, we can use Equation 15.4.26 and the method of images to find the Green's function for

$$\frac{\partial^2 g}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 g}{\partial t^2} = \delta(x-\xi)\delta(t-\tau), \quad 0 < x, t, \xi, \tau, \qquad \textbf{(15.4.31)}$$

subject to the boundary condition $g_x(0,t|\xi,\tau) = 0$.

[13] In electromagnetic theory, a free-space Green's function is the particular solution of the differential equation valid over a domain of infinite extent, where the Green's function remains bounded as we approach infinity, or satisfies a radiation condition there.

We begin by examining the related problem

$$\frac{\partial^2 g}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 g}{\partial t^2} = \delta(x-\xi)\delta(t-\tau) + \delta(x+\xi)\delta(t-\tau), \tag{15.4.32}$$

where $-\infty < x, \xi < \infty$, and $0 < t, \tau$. In this particular case, we have chosen an image that is the mirror reflection of $\delta(x-\xi)$. This was dictated by the fact that the Green's function must be an even function of x along $x = 0$ for any time t. In line with this argument,

$$g(x,t|\xi,\tau) = \frac{c}{2}\left[H\left(t-\tau-|x-\xi|/c\right) + H\left(t-\tau-|x+\xi|/c\right)\right]. \tag{15.4.33}$$

□

• Example 15.4.3: Equation of telegraphy

When the vibrating string problem includes the effect of air resistance, Equation 15.4.21 becomes

$$\frac{\partial^2 g}{\partial t^2} + 2\gamma\frac{\partial g}{\partial t} - c^2\frac{\partial^2 g}{\partial x^2} = c^2\delta(x-\xi)\delta(t-\tau), \tag{15.4.34}$$

where $-\infty < x, \xi < \infty$, and $0 < t, \tau$, with the boundary conditions

$$\lim_{|x|\to\infty} g(x,t|\xi,\tau) \to 0 \tag{15.4.35}$$

and the initial conditions

$$g(x,0|\xi,\tau) = g_t(x,0|\xi,\tau) = 0. \tag{15.4.36}$$

Let us find the Green's function.

Our analysis begins by introducing an intermediate dependent variable $w(x,t|\xi,\tau)$, where $g(x,t|\xi,\tau) = e^{-\gamma t}w(x,t|\xi,\tau)$. Substituting for $g(x,t|\xi,\tau)$, we now have

$$\frac{\partial^2 w}{\partial t^2} - \gamma^2 w - c^2\frac{\partial^2 w}{\partial x^2} = c^2\delta(x-\xi)\delta(t-\tau)e^{\gamma\tau}. \tag{15.4.37}$$

Taking the Laplace transform of Equation 15.4.37, we obtain

$$\frac{d^2 W}{dx^2} - \left(\frac{s^2-\gamma^2}{c^2}\right)W = -\delta(x-\xi)e^{\gamma\tau - s\tau}. \tag{15.4.38}$$

Using Fourier transforms as in Example 15.3.1, the solution to Equation 15.4.38 is

$$W(x,s|\xi,\tau) = \frac{\exp[-|x-\xi|\sqrt{(s^2-\gamma^2)/c^2} + \gamma\tau - s\tau]}{2\sqrt{(s^2-\gamma^2)/c^2}}. \tag{15.4.39}$$

Employing tables to invert the Laplace transform and the second shifting theorem, we have that

$$w(x,t|\xi,\tau) = \frac{c}{2}e^{\gamma\tau}I_0\left[\gamma\sqrt{(t-\tau)^2-(x-\xi)^2/c^2}\right]H[c(t-\tau)-|x-\xi|], \tag{15.4.40}$$

or

$$\boxed{g(x,t|\xi,\tau) = \frac{c}{2}e^{-\gamma(t-\tau)}I_0\left[\gamma\sqrt{(t-\tau)^2-(x-\xi)^2/c^2}\right]H[c(t-\tau)-|x-\xi|].} \tag{15.4.41}$$

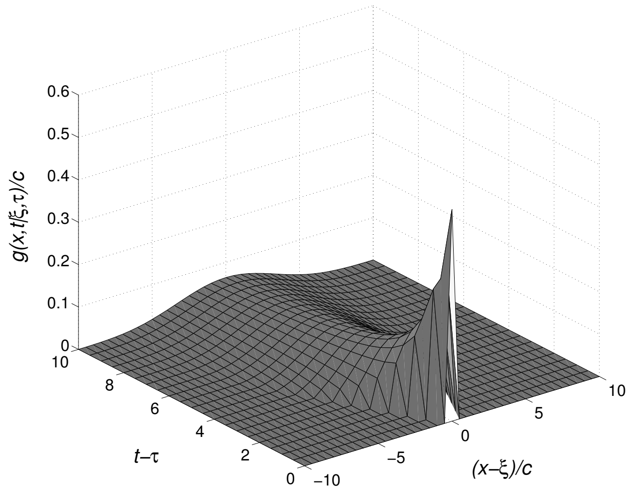


Figure 15.4.2: The free-space Green's function $g(x,t|\xi,\tau)/c$ for the one-dimensional equation of telegraphy with $\gamma = 1$ at different distances $(x-\xi)/c$ and times $t-\tau$.

Figure 15.4.2 illustrates Equation 15.4.41 when $\gamma = 1$. □

• Example 15.4.4

Let us solve[14] the one-dimensional wave equation on an infinite domain:

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = \cos(\omega t)\delta[x - X(t)], \tag{15.4.42}$$

subject to the boundary conditions

$$\lim_{|x|\to\infty} u(x,t) \to 0, \qquad 0 < t, \tag{15.4.43}$$

and initial conditions

$$u(x,0) = u_t(x,0) = 0, \qquad -\infty < x < \infty. \tag{15.4.44}$$

Here ω is a constant and $X(t)$ is some function of time.

With the given boundary and initial conditions, only the first integral in Equation 15.4.9 does not vanish. Substituting the source term $q(x,t) = \cos(\omega t)\delta[x - X(t)]$ and the Green's function given by Equation 15.4.26, we have that

$$u(x,t) = \int_0^t \int_{-\infty}^{\infty} q(\xi,\tau) g(x,t|\xi,\tau)\, d\xi\, d\tau \tag{15.4.45}$$

$$= \frac{1}{2}\int_0^t \int_{-\infty}^{\infty} \cos(\omega\tau)\delta[\xi - X(\tau)]H(t-\tau) - |x-\xi|)\, d\xi\, d\tau \tag{15.4.46}$$

$$= \frac{1}{2}\int_0^t H[t - \tau - |X(\tau) - x|]\cos(\omega\tau)\, d\tau, \tag{15.4.47}$$

[14] See Knowles, J. K., 1968: Propagation of one-dimensional waves from a source in random motion. *J. Acoust. Soc. Am.*, **43**, 948–957.

since $c = 1$.

Problems

1. By direct substitution, show[15] that

$$g(x,t|0,0) = J_0(\sqrt{xt}\,)H(x)H(t)$$

is the free-space Green's function governed by

$$\frac{\partial^2 g}{\partial x \partial t} + \tfrac{1}{4}g = \delta(x)\delta(t), \qquad -\infty < x, t < \infty.$$

2. Use Equation 15.3.20 to construct the Green's function for the one-dimensional wave equation

$$\frac{\partial^2 g}{\partial t^2} - \frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad 0 < x, \xi < L, \quad 0 < t, \tau,$$

subject to the boundary conditions $g(0,t|\xi,\tau) = g_x(L,t|\xi,\tau) = 0$, $0 < t$, and the initial conditions that $g(x,0|\xi,\tau) = g_t(x,0|\xi,\tau) = 0$, $0 < x < L$.

3. Use Equation 15.3.20 to construct the Green's function for the one-dimensional wave equation

$$\frac{\partial^2 g}{\partial t^2} - \frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad 0 < x, \xi < L, \quad 0 < t, \tau,$$

subject to the boundary conditions $g_x(0,t|\xi,\tau) = g_x(L,t|\xi,\tau) = 0$, $0 < t$, and the initial conditions that $g(x,0|\xi,\tau) = g_t(x,0|\xi,\tau) = 0$, $0 < x < L$.

4. Use the Green's function given by Equation 15.3.26 to write down the solution to the wave equation $u_{tt} = u_{xx}$ on the interval $0 < x < L$ with the boundary conditions $u(0,t) = u(L,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = \cos(\pi x/L)$ and $u_t(x,0) = 0$, $0 < x < L$.

5. Use the Green's function given by Equation 15.3.26 to write down the solution to the wave equation $u_{tt} = u_{xx}$ on the interval $0 < x < L$ with the boundary conditions $u(0,t) = e^{-t}$ and $u(L,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = \sin(\pi x/L)$ and $u_t(x,0) = 1$, $0 < x < L$.

6. Use the Green's function that you found in Problem 2 to write down the solution to the wave equation $u_{tt} = u_{xx}$ on the interval $0 < x < L$ with the boundary conditions $u(0,t) = 0$ and $u_x(L,t) = 1$, $0 < t$, and the initial conditions $u(x,0) = x$ and $u_t(x,0) = 1$, $0 < x < L$.

7. Use the Green's function that you found in Problem 3 to write down the solution to the wave equation $u_{tt} = u_{xx}$ on the interval $0 < x < L$ with the boundary conditions $u_x(0,t) = 1$ and $u_x(L,t) = 0$, $0 < t$, and the initial conditions $u(x,0) = 1$ and $u_t(x,0) = 0$, $0 < x < L$.

[15] First proven by Picard, É., 1894: Sur une équation aux dérivées partielles de la théorie de la propagation de l'électricité. *Bull. Soc. Math.*, **22**, 2–8.

8. Find the Green's function[16] governed by

$$\frac{\partial^2 g}{\partial t^2} + 2\frac{\partial g}{\partial t} - \frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad 0 < x, \xi < L, \quad 0 < t, \tau,$$

subject to the boundary conditions

$$g_x(0, t|\xi, \tau) = g_x(L, t|\xi, \tau) = 0, \qquad 0 < t,$$

and the initial conditions

$$g(x, 0|\xi, \tau) = g_t(x, 0|\xi, \tau) = 0, \qquad 0 < x < L.$$

Step 1: If the Green's function can be written as the Fourier half-range cosine series

$$g(x, t|\xi, \tau) = \frac{1}{L}G_0(t|\tau) + \frac{2}{L}\sum_{n=1}^{\infty} G_n(t|\tau)\cos\left(\frac{n\pi x}{L}\right),$$

so that it satisfies the boundary conditions, show that $G_n(t|\tau)$ is governed by

$$G_n'' + 2G_n' + \frac{n^2\pi^2}{L^2}G_n = \cos\left(\frac{n\pi\xi}{L}\right)\delta(t-\tau), \qquad 0 \le n.$$

Step 2: Show that

$$G_0(t|\tau) = e^{-(t-\tau)}\sinh(t-\tau)H(t-\tau),$$

and

$$G_n(t|\tau) = \cos\left(\frac{n\pi\xi}{L}\right)e^{-(t-\tau)}\frac{\sin[\beta_n(t-\tau)]}{\beta_n}H(t-\tau), \quad 1 \le n,$$

where $\beta_n = \sqrt{(n\pi/L)^2 - 1}$.

Step 3: Combine the results from Steps 1 and 2 and show that

$$\begin{aligned} g(x, t|\xi, \tau) = {} & e^{-(t-\tau)}\sinh(t-\tau)H(t-\tau)/L \\ & + 2e^{-(t-\tau)}H(t-\tau)/L \\ & \quad \times \sum_{n=1}^{\infty}\frac{\sin[\beta_n(t-\tau)]}{\beta_n}\cos\left(\frac{n\pi\xi}{L}\right)\cos\left(\frac{n\pi x}{L}\right). \end{aligned}$$

[16] Özişik, M. N., and B. Vick, 1984: Propagation and reflection of thermal waves in a finite medium. *Int. J. Heat Mass Transfer*, **27**, 1845–1854; Tang, D.-W., and N. Araki, 1996: Propagation of non-Fourier temperature wave in finite medium under laser-pulse heating (in Japanese). *Nihon Kikai Gakkai Rombumshu (Trans. Japan Soc. Mech. Engrs.), Ser. B*, **62**, 1136–1141.

15.5 HEAT EQUATION

In this section we present the Green's function[17] for the heat equation

$$\frac{\partial u}{\partial t} - a^2 \frac{\partial^2 u}{\partial x^2} = q(x, t), \tag{15.5.1}$$

where t denotes time, x is the position, a^2 is the diffusivity, and $q(x, t)$ is the source density. In addition to Equation 15.5.1, boundary conditions must be specified to ensure the uniqueness of solution; the most common ones are Dirichlet, Neumann, and Robin (a linear combination of the first two). An initial condition $u(x, t = t_0)$ is also needed.

The heat equation differs in many ways from the wave equation and the Green's function must, of course, manifest these differences. The most notable one is the asymmetry of the heat equation with respect to time. This merely reflects the fact that the heat equation differentiates between past and future as entropy continually increases.

We begin by proving that we can express the solution to Equation 15.5.1 in terms of boundary conditions, the initial condition, and the Green's function, which is found by solving

$$\frac{\partial g}{\partial t} - a^2 \frac{\partial^2 g}{\partial x^2} = \delta(x - \xi)\delta(t - \tau), \tag{15.5.2}$$

where ξ denotes the position of the source. From causality[18] we know that $g(x, t|\xi, \tau) = 0$ if $t < \tau$. We again require that the Green's function $g(x, t|\xi, \tau)$ satisfies the homogeneous form of the boundary condition on $u(x, t)$. For example, if u satisfies a homogeneous or nonhomogeneous Dirichlet condition, then the Green's function will satisfy the corresponding *homogeneous* Dirichlet condition. Although we will focus on the mathematical aspects of the problem, Equation 15.5.2 can be given the physical interpretation of the temperature distribution within a medium when a unit of heat is introduced at ξ at time τ.

We now establish that the solution to the nonhomogeneous heat equation can be expressed in terms of the Green's function, boundary conditions, and the initial condition. We begin with the equations

$$a^2 \frac{\partial^2 u(\xi, \tau)}{\partial \xi^2} - \frac{\partial u(\xi, \tau)}{\partial \tau} = -q(\xi, \tau), \tag{15.5.3}$$

and

$$a^2 \frac{\partial^2 g(x, t|\xi, \tau)}{\partial \xi^2} + \frac{\partial g(x, t|\xi, \tau)}{\partial \tau} = -\delta(x - \xi)\delta(t - \tau). \tag{15.5.4}$$

As we did in the previous section, we multiply Equation 15.5.3 by $g(x, t|\xi, \tau)$ and Equation 15.5.4 by $u(\xi, \tau)$ and subtract. Integrating over ξ from a to b, where a and b are the endpoints of the spatial domain, and over τ from 0 to t^+, where t^+ denotes a time slightly later than t so that we avoid ending the integration exactly at the peak of the delta function, we find

$$a^2 \int_0^{t^+} \int_a^b \left[u(\xi, \tau) \frac{\partial^2 g(x, t|\xi, \tau)}{\partial \xi^2} - g(x, t|\xi, \tau) \frac{\partial^2 u(\xi, \tau)}{\partial \xi^2} \right] d\xi \, d\tau$$

[17] See also Carslaw, H. S., and J. C. Jaeger, 1959: *Conduction of Heat in Solids*. Clarendon Press, Chapter 14; Beck, J. V., K. D. Cole, A. Haji-Sheikh, and B. Litkouhi, 1992: *Heat Conduction Using Green's Functions*. Hemisphere Publishing Corp., 523 pp.; Özişik, M. N., 1993: *Heat Conduction*. John Wiley & Sons, Inc., Chapter 6.

[18] The principle stating that an event cannot precede its cause.

$$+\int_0^{t^+}\int_a^b \left[u(\xi,\tau)\frac{\partial g(x,t|\xi,\tau)}{\partial \tau} + g(x,t|\xi,\tau)\frac{\partial u(\xi,\tau)}{\partial \tau}\right] d\xi\, d\tau$$
$$= \int_0^{t^+}\int_a^b q(\xi,\tau)g(x,t|\xi,\tau)\, d\xi\, d\tau - u(x,t). \qquad \textbf{(15.5.5)}$$

Applying Equation 15.4.6 and performing the time integration in the second integral, we finally obtain

$$u(x,t) = \int_0^{t^+}\int_a^b q(\xi,\tau)g(x,t|\xi,\tau)\, d\xi\, d\tau$$
$$+ a^2 \int_0^{t^+} \left[g(x,t|\xi,\tau)\frac{\partial u(\xi,\tau)}{\partial \xi} - u(\xi,\tau)\frac{\partial g(x,t|\xi,\tau)}{\partial \xi}\right]_{\xi=a}^{\xi=b} d\tau$$
$$+ \int_a^b u(\xi,0)g(x,t|\xi,0)\, d\xi, \qquad \textbf{(15.5.6)}$$

where we used $g(x,t|\xi,t^+) = 0$. The first two terms in Equation 15.5.6 represent the familiar effects of volume sources and boundary conditions, while the third term includes the effects of the initial data.

• Example 15.5.1: One-dimensional heat equation in an unlimited domain

The Green's function for the one-dimensional heat equation is governed by

$$\frac{\partial g}{\partial t} - a^2\frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad -\infty < x,\xi < \infty, \quad 0 < t,\tau, \qquad \textbf{(15.5.7)}$$

subject to the boundary conditions $\lim_{|x|\to\infty} g(x,t|\xi,\tau) \to 0$, and the initial condition $g(x,0|\xi,\tau) = 0$. Let us find $g(x,t|\xi,\tau)$.

We begin by taking the Laplace transform of Equation 15.5.7 and find that

$$\frac{d^2G}{dx^2} - \frac{s}{a^2}G = -\frac{\delta(x-\xi)}{a^2}e^{-s\tau}. \qquad \textbf{(15.5.8)}$$

Next, we take the Fourier transform of Equation 15.5.8 so that

$$(k^2+b^2)\overline{G}(k,s|\xi,\tau) = \frac{e^{-ik\xi}e^{-s\tau}}{a^2}, \qquad \textbf{(15.5.9)}$$

where $\overline{G}(k,s|\xi,\tau)$ is the Fourier transform of $G(x,s|\xi,\tau)$ and $b^2 = s/a^2$.

To find $G(x,s|\xi,\tau)$, we use the inversion integral

$$G(x,s|\xi,\tau) = \frac{e^{-s\tau}}{2\pi a^2}\int_{-\infty}^{\infty}\frac{e^{i(x-\xi)k}}{k^2+b^2}\, dk. \qquad \textbf{(15.5.10)}$$

Transforming Equation 15.5.10 into a closed contour via Jordan's lemma, we evaluate it by the residue theorem and find that

$$G(x,s|\xi,\tau) = \frac{e^{-|x-\xi|\sqrt{s}/a - s\tau}}{2a\sqrt{s}}. \qquad \textbf{(15.5.11)}$$

From a table of Laplace transforms we finally obtain

$$g(x,t|\xi,\tau) = \frac{H(t-\tau)}{\sqrt{4\pi a^2(t-\tau)}} \exp\left[-\frac{(x-\xi)^2}{4a^2(t-\tau)}\right], \qquad \textbf{(15.5.12)}$$

after applying the second shifting theorem. □

The primary use of the fundamental or free-space Green's function[19] is as a *particular* solution to the Green's function problem. For this reason, it is often called the *fundamental heat conduction solution*. Consequently, we usually must find a homogeneous solution so that the sum of the free-space Green's function plus the homogeneous solution satisfies any boundary conditions. The following examples show some commonly employed techniques.

• Example 15.5.2

Let us find the Green's function for the following problem:

$$\frac{\partial g}{\partial t} - a^2\frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad 0 < x, \xi < \infty, \quad 0 < t, \tau, \qquad \textbf{(15.5.13)}$$

subject to the boundary conditions $g(0,t|\xi,\tau) = 0$, $\lim_{x\to\infty} g(x,t|\xi,\tau) \to 0$, and the initial condition $g(x,0|\xi,\tau) = 0$. From the boundary condition $g(0,t|\xi,\tau) = 0$, we deduce that $g(x,t|\xi,\tau)$ must be an odd function in x over the open interval $(-\infty,\infty)$. We find this Green's function by introducing an image source of $-\delta(x+\xi)$ and resolving Equation 15.5.7 with the source $\delta(x-\xi)\delta(t-\tau) - \delta(x+\xi)\delta(t-\tau)$. Because Equation 15.5.7 is linear, Equation 15.5.12 gives the solution for each delta function and the Green's function for Equation 15.5.13 is

$$g(x,t|\xi,\tau) = \frac{H(t-\tau)}{\sqrt{4\pi a^2(t-\tau)}}\left\{\exp\left[-\frac{(x-\xi)^2}{4a^2(t-\tau)}\right] - \exp\left[-\frac{(x+\xi)^2}{4a^2(t-\tau)}\right]\right\} \qquad \textbf{(15.5.14)}$$

$$= \frac{H(t-\tau)}{\sqrt{\pi a^2(t-\tau)}} \exp\left[-\frac{x^2+\xi^2}{4a^2(t-\tau)}\right] \sinh\left[\frac{x\xi}{2a^2(t-\tau)}\right]. \qquad \textbf{(15.5.15)}$$

In a similar manner, if the boundary condition at $x = 0$ changes to $g_x(0,t|\xi,\tau) = 0$, then Equation 15.5.14 through Equation 15.5.15 become

$$g(x,t|\xi,\tau) = \frac{H(t-\tau)}{\sqrt{4\pi a^2(t-\tau)}}\left\{\exp\left[-\frac{(x-\xi)^2}{4a^2(t-\tau)}\right] + \exp\left[-\frac{(x+\xi)^2}{4a^2(t-\tau)}\right]\right\} \qquad \textbf{(15.5.16)}$$

$$= \frac{H(t-\tau)}{\sqrt{\pi a^2(t-\tau)}} \exp\left[-\frac{x^2+\xi^2}{4a^2(t-\tau)}\right] \cosh\left[\frac{x\xi}{2a^2(t-\tau)}\right]. \qquad \textbf{(15.5.17)}$$

---

[19] In electromagnetic theory, a free-space Green's function is the particular solution of the differential equation valid over a domain of infinite extent, where the Green's function remains bounded as we approach infinity, or satisfies a radiation condition there.

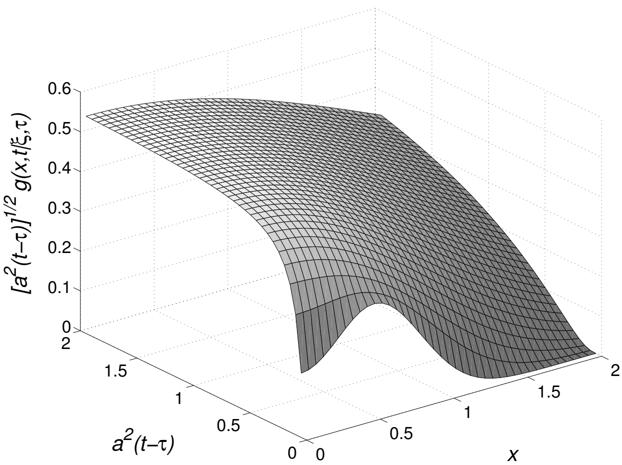


Figure 15.5.1: The Green's function, Equation 15.5.17, for the one-dimensional heat equation on the semi-infinite domain $0 < x < \infty$, and $0 \le t - \tau$, when the left boundary condition is $g_x(0,t|\xi,\tau) = 0$ and $\xi = 0.5$.

Figure 15.5.1 illustrates Equation 15.5.17 for the special case when $\xi = 0.5$. □

• Example 15.5.3: One-dimensional heat equation on the interval $0 < x < L$

Here we find the Green's function for the one-dimensional heat equation over the interval $0 < x < L$ associated with the problem

$$\frac{\partial u}{\partial t} - a^2 \frac{\partial^2 u}{\partial x^2} = f(x,t), \quad 0 < x < L, \quad 0 < t, \tag{15.5.18}$$

where a^2 is the diffusivity constant.

To find the Green's function for this problem, consider the following problem:

$$\frac{\partial g}{\partial t} - a^2 \frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad 0 < x, \xi < L, \quad 0 < t, \tau, \tag{15.5.19}$$

with the boundary conditions

$$\alpha_1 g(0,t|\xi,\tau) + \beta_1 g_x(0,t|\xi,\tau) = 0, \quad 0 < t, \tag{15.5.20}$$

and

$$\alpha_2 g(L,t|\xi,\tau) + \beta_2 g_x(L,t|\xi,\tau) = 0, \quad 0 < t, \tag{15.5.21}$$

and the initial condition

$$g(x,0|\xi,\tau) = 0, \quad 0 < x < L. \tag{15.5.22}$$

We begin by taking the Laplace transform of Equation 15.5.19 and find that

$$\frac{d^2 G}{dx^2} - \frac{s}{a^2} G = -\frac{\delta(x-\xi)}{a^2} e^{-s\tau}, \qquad 0 < x < L, \tag{15.5.23}$$

with

$$\alpha_1 G(0,s|\xi,\tau) + \beta_1 G'(0,s|\xi,\tau) = 0, \tag{15.5.24}$$

and

$$\alpha_2 G(L, s|\xi, \tau) + \beta_2 G'(L, s|\xi, \tau) = 0. \tag{15.5.25}$$

Problems similar to Equation 15.5.23 through Equation 15.5.25 were considered in Section 15.2. Applying this technique of eigenfunction expansions, we have that

$$G(x, s|\xi, \tau) = e^{-s\tau} \sum_{n=1}^{\infty} \frac{\varphi_n(\xi)\varphi_n(x)}{s + a^2 k_n^2}, \tag{15.5.26}$$

where $\varphi_n(x)$ is the nth *orthonormal* eigenfunction to the regular Sturm-Liouville problem

$$\varphi''(x) + k^2\varphi(x) = 0, \qquad 0 < x < L, \tag{15.5.27}$$

subject to the boundary conditions

$$\alpha_1\varphi(0) + \beta_1\varphi'(0) = 0, \tag{15.5.28}$$

and

$$\alpha_2\varphi(L) + \beta_2\varphi'(L) = 0. \tag{15.5.29}$$

Taking the inverse of Equation 15.5.26, we have that

$$g(x, t|\xi, \tau) = \left[\sum_{n=1}^{\infty} \varphi_n(\xi)\varphi_n(x)e^{-k_n^2 a^2 (t-\tau)}\right] H(t - \tau). \tag{15.5.30}$$

For example, let us find the Green's function for the heat equation on a finite domain

$$\frac{\partial g}{\partial t} - a^2\frac{\partial^2 g}{\partial x^2} = \delta(x - \xi)\delta(t - \tau), \quad 0 < x, \xi < L, \quad 0 < t, \tau, \tag{15.5.31}$$

with the boundary conditions $g(0, t|\xi, \tau) = g(L, t|\xi, \tau) = 0$, $0 < t$, and the initial condition $g(x, 0|\xi, \tau) = 0$, $0 < x < L$.

The Sturm-Liouville problem is

$$\varphi''(x) + k^2\varphi(x) = 0, \qquad 0 < x < L, \tag{15.5.32}$$

with the boundary conditions $\varphi(0) = \varphi(L) = 0$. The nth *orthonormal* eigenfunction to Equation 15.5.32 is

$$\varphi_n(x) = \sqrt{\frac{2}{L}}\,\sin\left(\frac{n\pi x}{L}\right). \tag{15.5.33}$$

Substituting Equation 15.5.33 into Equation 15.5.30, we find that

$$g(x, t|\xi, \tau) = \frac{2}{L}\left\{\sum_{n=1}^{\infty} \sin\left(\frac{n\pi\xi}{L}\right)\sin\left(\frac{n\pi x}{L}\right)e^{-a^2n^2\pi^2(t-\tau)/L^2}\right\} H(t - \tau). \tag{15.5.34}$$

On the other hand, the Green's function for the heat equation on a finite domain governed by

$$\frac{\partial g}{\partial t} - a^2 \frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \quad 0 < x, \xi < L, \quad 0 < t, \tau, \tag{15.5.35}$$

with the boundary conditions

$$g_x(0, t|\xi, \tau) = 0, \quad g_x(L, t|\xi, \tau) + hg(L, t|\xi, \tau) = 0, \quad 0 < t, \tag{15.5.36}$$

and the initial condition $g(x, 0|\xi, \tau) = 0$, $0 < x < L$, yields the Sturm-Liouville problem that we must now solve:

$$\varphi''(x) + \lambda\varphi(x) = 0, \quad \varphi'(0) = 0, \quad \varphi'(L) + h\varphi(L) = 0. \tag{15.5.37}$$

The nth *orthonormal* eigenfunction for Equation 15.5.37 is

$$\varphi_n(x) = \sqrt{\frac{2(k_n^2 + h^2)}{L(k_n^2 + h^2) + h}} \cos(k_n x), \tag{15.5.38}$$

where k_n is the nth root of $k\tan(kL) = h$. We also used the identity that $(k_n^2 + h^2)\sin^2(k_n h) = h^2$. Substituting Equation 15.5.38 into Equation 15.5.30, we finally obtain

$$g(x, t|\xi, \tau) = \frac{2}{L}\left\{\sum_{n=1}^{\infty} \frac{[(k_n L)^2 + (hL)^2]\cos(k_n\xi)\cos(k_n x)}{(k_n L)^2 + (hL)^2 + hL} e^{-a^2 k_n^2 (t-\tau)}\right\} H(t-\tau). \tag{15.5.39}$$

□

• Example 15.5.4

Let us use Green's functions to solve the heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad 0 < x < L, \quad 0 < t, \tag{15.5.40}$$

subject to the boundary conditions

$$u(0, t) = 0, \quad u(L, t) = t, \qquad 0 < t, \tag{15.5.41}$$

and the initial condition

$$u(x, 0) = 0, \qquad 0 < x < L. \tag{15.5.42}$$

Because there is no source term, Equation 15.5.6 simplifies to

$$\begin{aligned} u(x, t) = a^2 &\int_0^t [g(x, t|L, \tau)u_\xi(L, \tau) - u(L, \tau)g_\xi(x, t|L, \tau)]\, d\tau \\ &- a^2 \int_0^t [g(x, t|0, \tau)u_\xi(0, \tau) - u(0, \tau)g_\xi(x, t|0, \tau)]\, d\tau + \int_0^L u(\xi, 0)g(x, t|\xi, 0)\, d\xi. \end{aligned} \tag{15.5.43}$$

The Green's function that should be used here is the one given by Equation 15.5.34. Further simplification occurs by noting that $g(x, t|0, \tau) = g(x, t|L, \tau) = 0$ as well as $u(0, \tau) = u(\xi, 0) = 0$. Therefore we are left with the single integral

$$u(x, t) = -a^2 \int_0^t u(L, \tau)g_\xi(x, t|L, \tau)\, d\tau. \tag{15.5.44}$$

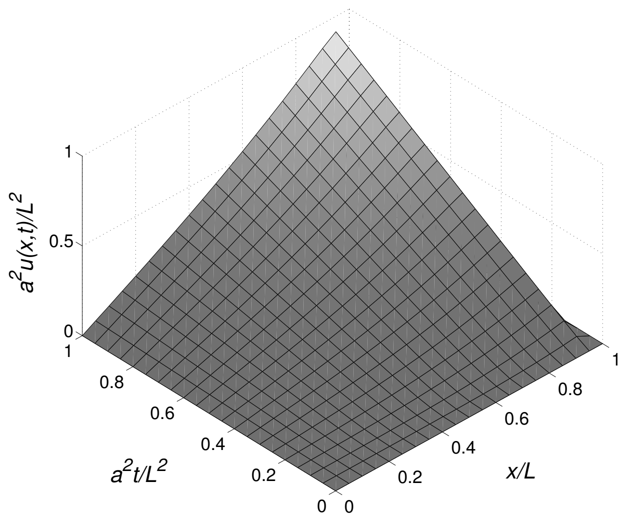


Figure 15.5.2: The temperature distribution within a bar when the temperature is initially at zero and then the ends are held at zero at $x = 0$ and t at $x = L$.

Upon substituting for $g(x,t|L,\tau)$ and reversing the order of integration and summation,

$$u(x,t) = -\frac{2\pi a^2}{L^2}\sum_{n=1}^{\infty}(-1)^n n \sin\left(\frac{n\pi x}{L}\right)\int_0^t \tau \exp\left[\frac{a^2n^2\pi^2}{L^2}(\tau - t)\right]\,d\tau \tag{15.5.45}$$

$$= -\frac{2L^2}{a^2\pi^3}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^3}\sin\left(\frac{n\pi x}{L}\right)\exp\left[\frac{a^2n^2\pi^2}{L^2}(\tau - t)\right]\left(\frac{a^2n^2\pi^2\tau}{L^2} - 1\right)\Bigg|_0^t \tag{15.5.46}$$

$$= -\frac{2L^2}{a^2\pi^3}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^3}\sin\left(\frac{n\pi x}{L}\right)\left[\frac{a^2n^2\pi^2 t}{L^2} - 1 + \exp\left(-\frac{a^2n^2\pi^2 t}{L^2}\right)\right]. \tag{15.5.47}$$

Figure 15.5.2 illustrates Equation 15.5.47. This solution is equivalent to Equation 8.6.16 that we found by Duhamel's integral. □

• Example 15.5.5: Heat equation within a cylinder

In this example, we find the Green's function for the heat equation in cylindrical coordinates

$$\frac{\partial g}{\partial t} - \frac{a^2}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}\right) = \frac{\delta(r-\rho)\delta(t-\tau)}{2\pi r}, \quad 0 < r, \rho < b, \quad 0 < t, \tau, \tag{15.5.48}$$

subject to the boundary conditions $\lim_{r\to 0}|g(r,t|\rho,\tau)| < \infty$, $g(b,t|\rho,\tau) = 0$, and the initial condition $g(r,0|\rho,\tau) = 0$.

As usual, we begin by taking the Laplace transform of Equation 15.5.48, or

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dG}{dr}\right) - \frac{s}{a^2}G = -\frac{e^{-s\tau}}{2\pi a^2 r}\delta(r-\rho). \tag{15.5.49}$$

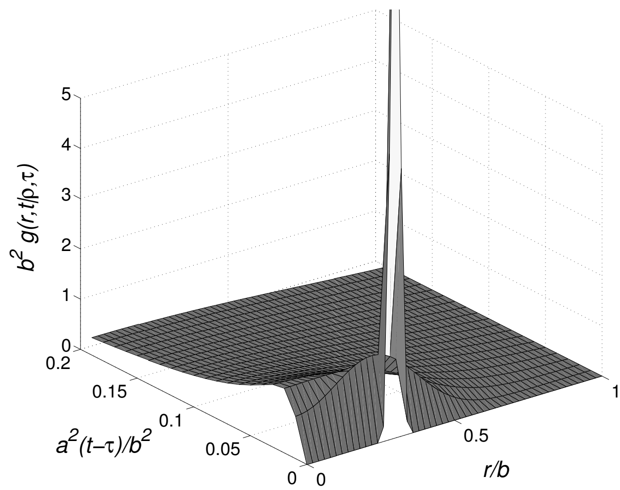


Figure 15.5.3: The Green's function, Equation 15.5.54, for the axisymmetric heat equation, Equation 15.5.48, with a Dirichlet boundary condition at $r = b$. Here $\rho/b = 0.3$ and the graph starts at $a^2(t-\tau)/b^2 = 0.001$ to avoid the delta function at $t - \tau = 0$.

Next we re-express $\delta(r-\rho)/r$ as the Fourier-Bessel expansion

$$\frac{\delta(r-\rho)}{2\pi r} = \sum_{n=1}^{\infty} A_n J_0(k_n r/b), \qquad \textbf{(15.5.50)}$$

where k_n is the nth root of $J_0(k) = 0$, and

$$A_n = \frac{2}{b^2 J_1^2(k_n)} \int_0^b \frac{\delta(r-\rho)}{2\pi r} J_0(k_n r/b)\, r\, dr = \frac{J_0(k_n \rho/b)}{\pi b^2 J_1^2(k_n)} \qquad \textbf{(15.5.51)}$$

so that

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dG}{dr}\right) - \frac{s}{a^2}G = -\frac{e^{-s\tau}}{\pi a^2 b^2} \sum_{n=1}^{\infty} \frac{J_0(k_n\rho/b) J_0(k_n r/b)}{J_1^2(k_n)}. \qquad \textbf{(15.5.52)}$$

The solution to Equation 15.5.52 is

$$G(r,s|\rho,\tau) = \frac{e^{-s\tau}}{\pi} \sum_{n=1}^{\infty} \frac{J_0(k_n\rho/b) J_0(k_n r/b)}{(sb^2 + a^2 k_n^2) J_1^2(k_n)}. \qquad \textbf{(15.5.53)}$$

Taking the inverse of Equation 15.5.53 and applying the second shifting theorem,

$$g(r,t|\rho,\tau) = \frac{H(t-\tau)}{\pi b^2} \sum_{n=1}^{\infty} \frac{J_0(k_n\rho/b) J_0(k_n r/b)}{J_1^2(k_n)} e^{-a^2 k_n^2 (t-\tau)/b^2}. \qquad \textbf{(15.5.54)}$$

See Figure 15.5.3.

If we modify the boundary condition at $r = b$ so that it now reads

$$g_r(b,t|\rho,\tau) + hg(b,t|\rho,\tau) = 0, \qquad \textbf{(15.5.55)}$$

where $h \geq 0$, our analysis now leads to

$$g(r,t|\rho,\tau) = \frac{H(t-\tau)}{\pi b^2} \sum_{n=1}^{\infty} \frac{J_0(k_n\rho/b) J_0(k_n r/b)}{J_0^2(k_n) + J_1^2(k_n)} e^{-a^2 k_n^2 (t-\tau)/b^2}, \qquad \textbf{(15.5.56)}$$

where k_n are the positive roots of $k\, J_1(k) - hb\, J_0(k) = 0$. If $h = 0$, we must add $1/(\pi b^2)$ to Equation 15.5.56.

Problems

1. Find the free-space Green's function for the linearized Ginzburg-Landau equation[20]

$$\frac{\partial g}{\partial t} + v\frac{\partial g}{\partial x} - ag - b\frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \ -\infty < x, \xi < \infty, \ 0 < t, \tau,$$

with $b > 0$.

Step 1: Taking the Laplace transform of the partial differential equation, show that it reduces to the ordinary differential equation

$$b\frac{d^2 G}{dx^2} - v\frac{dG}{dx} + aG - sG = -\delta(x-\xi)e^{-s\tau}.$$

Step 2: Using Fourier transforms, show that

$$G(x, s|\xi, \tau) = \frac{e^{-s\tau}}{2\pi}\int_{-\infty}^{\infty} \frac{e^{ik(x-\xi)}}{s + ikv + bk^2 - a}\, dk,$$

or

$$g(x, t|\xi, \tau) = \frac{e^{a(t-\tau)}}{\pi} H(t-\tau)\int_0^{\infty} e^{-b(t-\tau)k^2} \cos\{k[x - \xi - v(t-\tau)]\}\, dk.$$

Step 3: Evaluate the second integral and show that

$$g(x, t|\xi, \tau) = \frac{e^{a(t-\tau)} H(t-\tau)}{2\sqrt{\pi b(t-\tau)}} \exp\left\{-\frac{[x - \xi - v(t-\tau)]^2}{4b(t-\tau)}\right\}.$$

2. Use Green's functions to show that the solution[21] to

$$\frac{\partial u}{\partial t} = a^2\frac{\partial^2 u}{\partial x^2}, \qquad 0 < x, t,$$

subject to the boundary conditions

$$u(0, t) = 0, \quad \lim_{x\to\infty} u(x, t) \to 0, \qquad 0 < t,$$

and the initial condition

$$u(x, 0) = f(x), \qquad 0 < x < \infty,$$

is

$$u(x, t) = \frac{e^{-x^2/(4a^2 t)}}{a\sqrt{\pi t}} \int_0^{\infty} f(\tau) \sinh\left(\frac{x\tau}{2a^2 t}\right) e^{-\tau^2/(4a^2 t)}\, d\tau.$$

[20] See Deissler, R. J., 1985: Noise-sustained structure, intermittency, and the Ginzburg-Landau equation. *J. Stat. Phys.*, **40**, 371–395.

[21] See Gilev, S. D., and T. Yu. Mikhaĭlova, 1996: Current wave in shock compression of matter in a magnetic field. *Tech. Phys.*, **41**, 407–411.

3. Use Equation 15.5.30 to construct the Green's function for the one-dimensional heat equation $g_t - g_{xx} = \delta(x - \xi)\delta(t - \tau)$ for $0 < x < L$, $0 < t$, with the initial condition that $g(x, 0|\xi, \tau) = 0$, $0 < x < L$, and the boundary conditions that $g(0, t|\xi, \tau) = g_x(L, t|\xi, \tau) = 0$ for $0 < t$. Assume that $L \neq \pi$.

4. Use Equation 15.5.30 to construct the Green's function for the one-dimensional heat equation $g_t - g_{xx} = \delta(x - \xi)\delta(t - \tau)$ for $0 < x < L$, $0 < t$, with the initial condition that $g(x, 0|\xi, \tau) = 0$, $0 < x < L$, and the boundary conditions that $g_x(0, t|\xi, \tau) = g_x(L, t|\xi, \tau) = 0$ for $0 < t$.

5. Use Equation 15.5.43 and the Green's function given by Equation 15.5.34 to find the solution to the heat equation $u_t = u_{xx}$ for $0 < x < L$, $0 < t$, with the initial data $u(x, 0) = 1$, $0 < x < L$, and the boundary conditions $u(0, t) = e^{-t}$ and $u(L, t) = 0$ when $0 < t$.

6. Use Equation 15.5.43 and the Green's function that you found in Problem 3 to find the solution to the heat equation $u_t = u_{xx}$ for $0 < x < L$, $0 < t$, with the initial data $u(x, 0) = 1$, $0 < x < L$, and the boundary conditions $u(0, t) = \sin(t)$ and $u_x(L, t) = 0$ when $0 < t$.

7. Use Equation 15.5.43 and the Green's function that you found in Problem 4 to find the solution to the heat equation $u_t = u_{xx}$ for $0 < x < L$, $0 < t$, with the initial data $u(x, 0) = 1$, $0 < x < L$, and the boundary conditions $u_x(0, t) = 1$ and $u_x(L, t) = 0$ when $0 < t$.

8. Find the Green's function for

$$\frac{\partial g}{\partial t} - a^2 \frac{\partial^2 g}{\partial x^2} + a^2 k^2 g = \delta(x - \xi)\delta(t - \tau), \quad 0 < x, \xi < L, \quad 0 < t, \tau,$$

subject to the boundary conditions

$$g(0, t|\xi, \tau) = g_x(L, t|\xi, \tau) = 0, \qquad 0 < t,$$

and the initial condition

$$g(x, 0|\xi, \tau) = 0, \qquad 0 < x < L,$$

where a and k are real constants.

15.6 HELMHOLTZ'S EQUATION

In the previous sections, we sought solutions to the heat and wave equations via Green's functions. In this section, we turn to the reduced wave equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \lambda u = -f(x, y). \qquad \textbf{(15.6.1)}$$

Equation 15.6.1, generally known as *Helmholtz's equation*, includes the special case of *Poisson's equation* when $\lambda = 0$. Poisson's equation has a special place in the theory of Green's functions because George Green (1793–1841) invented his technique for its solution.

The reduced wave equation arises during the solution of the harmonically forced wave equation[22] by separation of variables. In one spatial dimension, the problem is

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = -f(x)e^{-i\omega t}. \tag{15.6.2}$$

Equation 15.6.2 occurs, for example, in the mathematical analysis of a stretched string over some interval subject to an external, harmonic forcing. Assuming that $u(x,t)$ is bounded everywhere, we seek solutions of the form $u(x,t) = y(x)e^{-i\omega t}$. Upon substituting this solution into Equation 15.6.2 we obtain the ordinary differential equation

$$y'' + k_0^2 y = -f(x), \tag{15.6.3}$$

where $k_0^2 = \omega^2/c^2$. This is an example of the one-dimensional Helmholtz equation.

Let us now use Green's functions to solve the Helmholtz equation, Equation 15.6.1, where the Green's function is given by the Helmholtz equation

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} + \lambda g = -\delta(x-\xi)\delta(y-\eta). \tag{15.6.4}$$

The most commonly encountered boundary conditions are

- the *Dirichlet boundary condition*, where g vanishes on the boundary,
- the *Neumann boundary condition*, where the normal gradient of g vanishes on the boundary, and
- the *Robin boundary condition*, which is the linear combination of the Dirichlet and Neumann conditions.

We begin by multiplying Equation 15.6.1 by $g(x,y|\xi,\eta)$ and Equation 15.6.4 by $u(x,y)$, subtract and integrate over the region $a < x < b, c < y < d$. We find that

$$\begin{aligned} u(\xi,\eta) = \int_c^d \int_a^b \biggl\{ g(x,y|\xi,\eta)\left[\frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2}\right] \\ - u(x,y)\left[\frac{\partial^2 g(x,y|\xi,\eta)}{\partial x^2} + \frac{\partial^2 g(x,y|\xi,\eta)}{\partial y^2}\right]\biggr\}\,dx\,dy \\ + \int_c^d \int_a^b f(x,y) g(x,y|\xi,\eta)\,dx\,dy \end{aligned} \tag{15.6.5}$$

$$\begin{aligned} = \int_c^d \int_a^b \left\{\frac{\partial}{\partial x}\left[g(x,y|\xi,\eta)\frac{\partial u(x,y)}{\partial x}\right] - \frac{\partial}{\partial x}\left[u(x,y)\frac{\partial g(x,y|\xi,\eta)}{\partial x}\right]\right\}\,dx\,dy \\ + \int_c^d \int_a^b \left\{\frac{\partial}{\partial y}\left[g(x,y|\xi,\eta)\frac{\partial u(x,y)}{\partial y}\right] - \frac{\partial}{\partial y}\left[u(x,y)\frac{\partial g(x,y|\xi,\eta)}{\partial y}\right]\right\}\,dx\,dy \\ + \int_c^d \int_a^b f(x,y) g(x,y|\xi,\eta)\,dx\,dy \end{aligned} \tag{15.6.6}$$

$$\begin{aligned} = \int_c^d \left[g(x,y|\xi,\eta)\frac{\partial u(x,y)}{\partial x} - u(x,y)\frac{\partial g(x,y|\xi,\eta)}{\partial x}\right]_{x=a}^{x=b} dy \\ + \int_a^b \left[g(x,y|\xi,\eta)\frac{\partial u(x,y)}{\partial y} - u(x,y)\frac{\partial g(x,y|\xi,\eta)}{\partial y}\right]_{y=c}^{y=d} dx \\ + \int_c^d \int_a^b f(x,y) g(x,y|\xi,\eta)\,dx\,dy. \end{aligned} \tag{15.6.7}$$

[22] See, for example, Graff, K. F., 1991: *Wave Motion in Elastic Solids*. Dover Publications, Inc., Section 1.4.

Because (ξ, η) is an arbitrary point inside the rectangle, we denote it in general by (x, y). Furthermore, the variable (x, y) is now merely a dummy integration variable that we now denote by (ξ, η). Upon making these substitutions and using the symmetry condition $g(x, y|\xi, \eta) = g(\xi, \eta|x, y)$, we have that

$$\begin{aligned} u(x, y) = & \int_c^d \left[g(x, y|\xi, \eta) \frac{\partial u(\xi, \eta)}{\partial \xi} - u(\xi, \eta) \frac{\partial g(x, y|\xi, \eta)}{\partial \xi} \right]_{\xi=a}^{\xi=b} d\eta \\ & + \int_a^b \left[g(x, y|\xi, \eta) \frac{\partial u(\xi, \eta)}{\partial \eta} - u(\xi, \eta) \frac{\partial g(x, y|\xi, \eta)}{\partial \eta} \right]_{\eta=c}^{\eta=d} d\xi \\ & + \int_c^d \int_a^b f(\xi, \eta) g(x, y|\xi, \eta) \, d\xi \, d\eta. \end{aligned} \qquad \mathbf{(15.6.8)}$$

Equation 15.6.8 shows that the solution of Helmholtz's equation depends upon the sources inside the rectangle and values of $u(x, y)$ and $(\partial u/\partial x, \partial u/\partial y)$ along the boundary. On the other hand, we must still find the particular Green's function for a given problem; this Green's function depends directly upon the boundary conditions. At this point, we work out several special cases.

1. *Nonhomogeneous Helmholtz equation and homogeneous Dirichlet boundary conditions*

In this case, let us assume that we can find a Green's function that also satisfies the same Dirichlet boundary conditions as $u(x, y)$. Once the Green's function is found, then Equation 15.6.8 reduces to

$$u(x, y) = \int_c^d \int_a^b f(\xi, \eta) g(x, y|\xi, \eta) \, d\xi d\eta. \qquad \mathbf{(15.6.9)}$$

A possible source of difficulty would be the nonexistence of the Green's function. From our experience in Section 15.2, we know that this will occur if λ equals one of the eigenvalues of the corresponding homogeneous problem. An example of this occurs in acoustics when the Green's function for the Helmholtz equation does not exist at *resonance*.

2. *Homogeneous Helmholtz equation and nonhomogeneous Dirichlet boundary conditions*

In this particular case, $f(x, y) = 0$. For convenience, let us use the Green's function from the previous example so that $g(x, y|\xi, \eta) = 0$ along all of the boundaries. Under these conditions, Equation 15.6.8 becomes

$$u(x, y) = -\int_a^b u(\xi, \eta) \frac{\partial g(x, y|\xi, \eta)}{\partial \eta} \bigg|_{\eta=c}^{\eta=d} d\xi - \int_c^d u(\xi, \eta) \frac{\partial g(x, y|\xi, \eta)}{\partial \xi} \bigg|_{\xi=a}^{\xi=b} d\eta. \qquad \mathbf{(15.6.10)}$$

Consequently, the solution is determined once we compute the normal gradient of the Green's function along the boundary.

3. *Nonhomogeneous Helmholtz equation and homogeneous Neumann boundary conditions*

If we require that $u(x,y)$ satisfies the nonhomogeneous Helmholtz equation with homogeneous Neumann boundary conditions, then the governing equations are Equation 15.6.1 and the boundary conditions $u_x = 0$ along $x = a$ and $x = b$, and $u_y = 0$ along $y = c$ and $y = d$. Integrating Equation 15.6.1, we have that

$$\int_c^d \left[\frac{\partial u(b,y)}{\partial x} - \frac{\partial u(a,y)}{\partial x}\right] dy + \int_a^b \left[\frac{\partial u(x,d)}{\partial y} - \frac{\partial u(x,c)}{\partial y}\right] dx + \lambda \int_c^d \int_a^b u(x,y)\, dx\, dy = -\int_c^d \int_a^b f(x,y)\, dx\, dy. \quad \mathbf{(15.6.11)}$$

Because the first two integrals in Equation 15.6.11 must vanish in the case of homogeneous Neumann boundary conditions, this equation cannot be satisfied if $\lambda = 0$ unless

$$\int_c^d \int_a^b f(x,y)\, dx\, dy = 0. \quad \mathbf{(15.6.12)}$$

A physical interpretation of Equation 15.6.12 is as follows: Consider the physical process of steady-state heat conduction within a rectangular region. The temperature $u(x,y)$ is given by Poisson's equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -f(x,y), \quad \mathbf{(15.6.13)}$$

where $f(x,y)$ is proportional to the density of the heat sources and sinks. The boundary conditions $u_x(a,y) = u_x(b,y) = 0$ and $u_y(x,c) = u_y(x,d) = 0$ imply that there is no heat exchange across the boundary. Consequently, no steady-state temperature distribution can exist unless the heat sources are balanced by heat sinks. This balance of heat sources and sinks is given by Equation 15.6.12.

Having provided an overview of how Green's functions can be used to solve Poisson and Helmholtz equations, let us now determine several of them for commonly encountered domains.

• Example 15.6.1: Free-space Green's function for the one-dimensional Helmholtz equation

Let us find the Green's function for the one-dimensional Helmholtz equation

$$g'' + k_0^2 g = -\delta(x-\xi), \qquad -\infty < x, \xi < \infty. \quad \mathbf{(15.6.14)}$$

If we solve Equation 15.6.14 by piecing together homogeneous solutions, then

$$g(x|\xi) = Ae^{-ik_0(x-\xi)} + Be^{ik_0(x-\xi)}, \quad \mathbf{(15.6.15)}$$

for $x < \xi$, while

$$g(x|\xi) = Ce^{-ik_0(x-\xi)} + De^{ik_0(x-\xi)}, \quad \mathbf{(15.6.16)}$$

for $\xi < x$.

Let us examine Equation 15.6.15 more closely. The solution represents two propagating waves. Because $x < \xi$, the first term is a wave propagating out to infinity, while the second term gives a wave propagating in from infinity. This is seen most clearly by including the $e^{-i\omega t}$ term into Equation 15.6.15, or

$$g(x|\xi)e^{-i\omega t} = Ae^{-ik_0(x-\xi)-i\omega t} + Be^{ik_0(x-\xi)-i\omega t}. \quad \mathbf{(15.6.17)}$$

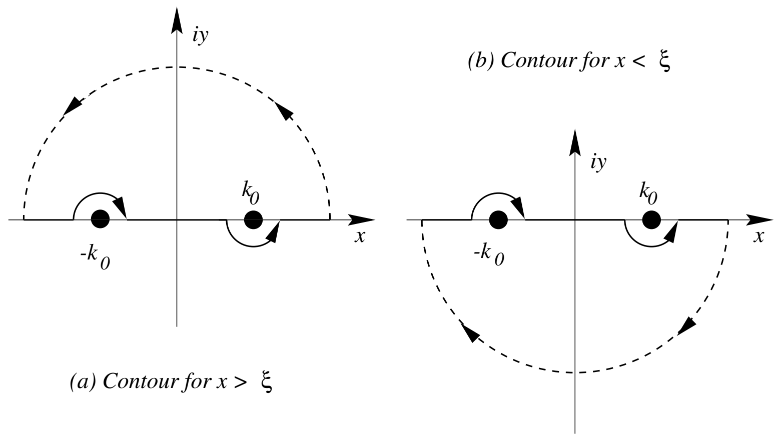


Figure 15.6.1: Contour used to evaluate Equation 15.6.21.

Because we have a source only at $x = \xi$, solutions that represent waves originating at infinity are nonphysical and we must discard them. This requirement that there are only outwardly propagating wave solutions is commonly called *Sommerfeld's radiation condition*.[23] Similar considerations hold for Equation 15.6.16 and we must take $C = 0$.

To evaluate A and D, we use the continuity conditions on the Green's function:

$$g(\xi^+|\xi) = g(\xi^-|\xi), \quad \text{and} \quad g'(\xi^+|\xi) - g'(\xi^-|\xi) = -1, \tag{15.6.18}$$

or

$$A = D, \qquad \text{and} \qquad ik_0 D + ik_0 A = -1. \tag{15.6.19}$$

Therefore,

$$\boxed{g(x|\xi) = \frac{i}{2k_0} e^{ik_0|x-\xi|}.} \tag{15.6.20}$$

We can also solve Equation 15.6.14 by Fourier transforms. Assuming that the Fourier transform of $g(x|\xi)$ exists and denoting it by $G(k|\xi)$, we find that

$$G(k|\xi) = \frac{e^{-ik\xi}}{k^2 - k_0^2}, \qquad \text{and} \qquad g(x|\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ik(x-\xi)}}{k^2 - k_0^2}\, dk. \tag{15.6.21}$$

Immediately we see that there is a problem with the singularities lying on the path of integration at $k = \pm k_0$. How do we avoid them?

There are four possible ways that we might circumvent the singularities. One of them is shown in Figure 15.6.1. Applying Jordan's lemma to close the line integral along the real

[23] Sommerfeld, A., 1912: Die Greensche Funktion der Schwingungsgleichung. *Jahresber. Deutschen Math.- Vereinung*, **21**, 309–353.

Free-Space Green's Function for the Poisson and Helmholtz Equations

Dimension	Poisson Equation	Helmholtz Equation
One	no solution	$g(x\|\xi) = \dfrac{i}{2k_0} e^{ik_0\|x-\xi\|}$
Two	$g(x,y\|\xi,\eta) = -\dfrac{\ln(r)}{2\pi}$	$g(x,y\|\xi,\eta) = \dfrac{i}{4} H_0^{(1)}(k_0 r)$
	$r = \sqrt{(x-\xi)^2 + (y-\eta)^2}$	

Note: For the Helmholtz equation, we have taken the temporal forcing to be $e^{-i\omega t}$ and $k_0 = \omega/c$.

axis (as shown in Figure 15.6.1),

$$g(x|\xi) = \frac{1}{2\pi} \oint_C \frac{e^{iz(x-\xi)}}{z^2 - k_0^2}\, dz. \tag{15.6.22}$$

For $x < \xi$,

$$g(x|\xi) = -i \text{ Res}\left[\frac{e^{iz(x-\xi)}}{z^2 - k_0^2}; -k_0\right] = \frac{i}{2k_0} e^{-ik_0(x-\xi)}, \tag{15.6.23}$$

while

$$g(x|\xi) = i \text{ Res}\left[\frac{e^{iz(x-\xi)}}{z^2 - k_0^2}; k_0\right] = \frac{i}{2k_0} e^{ik_0(x-\xi)}, \tag{15.6.24}$$

for $x > \xi$. A quick check shows that these solutions agree with Equation 15.6.20. If we try the three other possible paths around the singularities, we obtain incorrect solutions. □

• Example 15.6.2: Free-space Green's function for the two-dimensional Helmholtz equation

At this point, we have found two forms of the free-space Green's function for the one-dimensional Helmholtz equation. The first form is the analytic solution, Equation 15.6.20, while the second is the integral representation, Equation 15.6.21, where the line integration along the real axis is shown in Figure 15.6.1.

In the case of two dimensions, the Green's function[24] for the Helmholtz equation symmetric about the point (ξ, η) is the solution of the equation

$$\frac{d^2g}{dr^2} + \frac{1}{r}\frac{dg}{dr} + k_0^2 g = -\frac{\delta(r)}{2\pi r}, \tag{15.6.25}$$

where $r = \sqrt{(x-\xi)^2 + (y-\eta)^2}$. The homogeneous form of Equation 15.6.25 is Bessel's differential equation of order zero. Consequently, the general solution in terms of Hankel functions is

$$g(\mathbf{r}|\mathbf{r}_0) = A\, H_0^{(1)}(k_0 r) + B\, H_0^{(2)}(k_0 r). \tag{15.6.26}$$

[24] For an alternative derivation, see Graff, K. F., 1991: *Wave Motion in Elastic Solids*. Dover Publications, Inc., pp. 284–285.

Why have we chosen to use Hankel functions rather than $J_0(\cdot)$ and $Y_0(\cdot)$? As we argued earlier, solutions to the Helmholtz equation must represent *outwardly* propagating waves (the Sommerfeld radiation condition). If we again assume that the temporal behavior is $e^{-i\omega t}$ and use the asymptotic expressions for Hankel functions, we see that $H_0^{(1)}(k_0 r)$ represents outwardly propagating waves and $B = 0$.

What is the value of A? Integrating Equation 15.6.26 over a small circle around the point $r = 0$ and taking the limit as the radius of the circle vanishes, $A = i/4$ and

$$g(\mathbf{r}|\mathbf{r}_0) = \frac{i}{4} H_0^{(1)}(k_0 r). \tag{15.6.27}$$

If a real function is needed, then the free-space Green's function equals the Neumann function $Y_0(k_0 r)$ divided by -4. □

• Example 15.6.3: Free-space Green's function for the two-dimensional Laplace equation

In this subsection, we find the free-space Green's function for Poisson's equation in two dimensions. This Green's function is governed by

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} = -\frac{\delta(r-\rho)\delta(\theta-\theta')}{r}. \tag{15.6.28}$$

If we now choose our coordinate system so that the origin is located at the point source, $r = \sqrt{(x-\xi)^2 + (y-\eta)^2}$ and $\rho = 0$. Multiplying both sides of this simplified Equation 15.6.28 by $r\,dr\,d\theta$ and integrating over a circle of radius ϵ, we obtain -1 on the right side from the surface integration over the delta functions. On the left side,

$$\int_0^{2\pi} r\frac{\partial g}{\partial r}\bigg|_{r=\epsilon} d\theta = -1. \tag{15.6.29}$$

The Green's function $g(r,\theta|0,\theta') = -\ln(r)/(2\pi)$ satisfies Equation 15.6.29.

To find an alternative form of the free-space Green's function when the point of excitation and the origin of the coordinate system do not coincide, we first note that

$$\delta(\theta-\theta') = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{in(\theta-\theta')}. \tag{15.6.30}$$

This suggests that the Green's function should be of the form

$$g(r,\theta|\rho,\theta') = \sum_{n=-\infty}^{\infty} g_n(r|\rho)e^{in(\theta-\theta')}. \tag{15.6.31}$$

Substituting Equation 15.6.30 and Equation 15.6.31 into Equation 15.6.29, we obtain the ordinary differential equation

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dg_n}{dr}\right) - \frac{n^2}{r^2}g_n = -\frac{\delta(r-\rho)}{2\pi r}. \tag{15.6.32}$$

The homogeneous solution to Equation 15.6.32 is

$$g_0(r|\rho) = \begin{cases} a, & 0 \le r \le \rho\,, \\ b\,\ln(r), & \rho \le r < \infty. \end{cases} \tag{15.6.33}$$

and

$$g_n(r|\rho) = \begin{cases} c\,(r/\rho)^n, & 0 \le r \le \rho\,, \\ d\,(\rho/r)^n, & \rho \le r < \infty, \end{cases} \tag{15.6.34}$$

if $n \neq 0$.

At $r = \rho$, the g_n's must be continuous, in which case,

$$a = b\ln(\rho), \qquad \text{and} \qquad c = d. \tag{15.6.35}$$

On the other hand,

$$\rho \frac{dg_n}{dr}\bigg|_{r=\rho^-}^{r=\rho^+} = -\frac{1}{2\pi}, \tag{15.6.36}$$

or

$$a = -\frac{\ln(\rho)}{2\pi}, \quad b = -\frac{1}{2\pi}, \quad \text{and} \quad c = d = \frac{1}{4\pi n}. \tag{15.6.37}$$

Therefore,

$$g(r,\theta|\rho,\theta') = -\frac{\ln(r_>)}{2\pi} + \frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r_<}{r_>}\right)^n \cos[n(\theta-\theta')], \tag{15.6.38}$$

where $r_> = \max(r,\rho)$ and $r_< = \min(r,\rho)$.

We can simplify Equation 15.6.38 by noting that

$$\ln\left[1+\rho^2 - 2\rho\cos(\theta-\theta')\right] = -2\sum_{n=1}^{\infty}\frac{\rho^n\cos[n(\theta-\theta')]}{n}, \tag{15.6.39}$$

if $|\rho| < 1$. Applying this relationship to Equation 15.6.38, we find that

$$g(r,\theta|\rho,\theta') = -\frac{1}{4\pi}\ln\left[r^2+\rho^2-2r\rho\cos(\theta-\theta')\right]. \tag{15.6.40}$$

Note that when $\rho = 0$ we recover $g(r,\theta|0,\theta') = -\ln(r)/(2\pi)$. □

- **Example 15.6.4: Two-dimensional Poisson equation over a rectangular domain**

Consider the two-dimensional Poisson equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -f(x,y). \tag{15.6.41}$$

This equation arises in equilibrium problems, such as the static deflection of a rectangular membrane. In that case, $f(x,y)$ represents the external load per unit area, divided by the tension in the membrane. The solution $u(x,y)$ must satisfy certain boundary conditions. For the present, let us choose $u(0,y) = u(a,y) = 0$, and $u(x,0) = u(x,b) = 0$.

To find the Green's function for Equation 15.6.41 we must solve the partial differential equation

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta), \quad 0 < x, \xi < a, \quad 0 < y, \eta < b, \tag{15.6.42}$$

subject to the boundary conditions

$$g(0, y|\xi, \eta) = g(a, y|\xi, \eta) = g(x, 0|\xi, \eta) = g(x, b|\xi, \eta) = 0. \tag{15.6.43}$$

From Equation 15.6.9,

$$u(x, y) = \int_0^a \int_0^b g(x, y|\xi, \eta) f(\xi, \eta) \, d\eta \, d\xi. \tag{15.6.44}$$

One approach to finding the Green's function is to expand it in terms of the eigenfunctions $\varphi(x, y)$ of the differential equation

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} = -\lambda\varphi, \tag{15.6.45}$$

and the boundary conditions, Equation 15.6.43. The eigenvalues are

$$\lambda_{nm} = \frac{n^2\pi^2}{a^2} + \frac{m^2\pi^2}{b^2}, \tag{15.6.46}$$

where $n = 1, 2, 3, \ldots$, $m = 1, 2, 3, \ldots$, and the corresponding eigenfunctions are

$$\varphi_{nm}(x, y) = \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right). \tag{15.6.47}$$

Therefore, we seek $g(x, y|\xi, \eta)$ in the form

$$g(x, y|\xi, \eta) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{nm} \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right). \tag{15.6.48}$$

Because the delta functions can be written

$$\delta(x-\xi)\delta(y-\eta) = \frac{4}{ab} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \sin\left(\frac{n\pi\xi}{a}\right) \sin\left(\frac{m\pi\eta}{b}\right) \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right), \tag{15.6.49}$$

we find that

$$\left(\frac{n^2\pi^2}{a^2} + \frac{m^2\pi^2}{b^2}\right) A_{nm} = \frac{4}{ab} \sin\left(\frac{n\pi\xi}{a}\right) \sin\left(\frac{m\pi\eta}{b}\right), \tag{15.6.50}$$

after substituting Equations 15.6.48 and 15.6.49 into Equation 15.6.42, and setting the corresponding harmonics equal to each other. Therefore, the *bilinear formula* for the Green's function of Poisson's equation is

$$g(x, y|\xi, \eta) = \frac{4}{ab} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{\sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{n\pi\xi}{a}\right) \sin\left(\frac{m\pi y}{b}\right) \sin\left(\frac{m\pi\eta}{b}\right)}{n^2\pi^2/a^2 + m^2\pi^2/b^2}. \tag{15.6.51}$$

Thus, solutions to Poisson's equation can now be written as

$$u(x,y) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{a_{nm}}{n^2\pi^2/a^2 + m^2\pi^2/b^2} \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right), \tag{15.6.52}$$

where a_{nm} are the Fourier coefficients for the function $f(x,y)$:

$$a_{nm} = \frac{4}{ab}\int_0^a \int_0^b f(x,y) \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right)\, dy\, dx. \tag{15.6.53}$$

Another form of the Green's function can be obtained by considering each direction separately. To satisfy the boundary conditions along the edges $y = 0$ and $y = b$, we write the Green's function as the Fourier series

$$g(x,y|\xi,\eta) = \sum_{m=1}^{\infty} G_m(x|\xi) \sin\left(\frac{m\pi y}{b}\right), \tag{15.6.54}$$

where the coefficients $G_m(x|\xi)$ are left as undetermined functions of x, ξ, and m. Substituting this series into the partial differential equation for g, multiplying by $2\sin(n\pi y/b)/b$, and integrating over y, we find that

$$\frac{d^2G_n}{dx^2} - \frac{n^2\pi^2}{b^2}G_n = -\frac{2}{b}\sin\left(\frac{n\pi\eta}{b}\right)\delta(x-\xi). \tag{15.6.55}$$

This differential equation shows that the expansion coefficients $G_n(x|\xi)$ are one-dimensional Green's functions; we can find them, as we did in Section 15.2, by piecing together homogeneous solutions to Equation 15.6.55 that are valid over various intervals. For the region $0 \le x \le \xi$, the solution to this equation that vanishes at $x = 0$ is

$$G_n(x|\xi) = A_n \sinh\left(\frac{n\pi x}{b}\right), \tag{15.6.56}$$

where A_n is presently arbitrary. The corresponding solution for $\xi \le x \le a$ is

$$G_n(x|\xi) = B_n \sinh\left[\frac{n\pi(a-x)}{b}\right]. \tag{15.6.57}$$

Note that this solution vanishes at $x = a$. Because the Green's function must be continuous at $x = \xi$,

$$A_n \sinh\left(\frac{n\pi\xi}{b}\right) = B_n \sinh\left[\frac{n\pi(a-\xi)}{b}\right]. \tag{15.6.58}$$

On the other hand, the appropriate jump discontinuity of $G_n'(x|\xi)$ yields

$$-\frac{n\pi}{b}B_n \cosh\left[\frac{n\pi(a-\xi)}{b}\right] - \frac{n\pi}{b}A_n\cosh\left(\frac{n\pi\xi}{b}\right) = -\frac{2}{b}\sin\left(\frac{n\pi\eta}{b}\right). \tag{15.6.59}$$

Solving for A_n and B_n,

$$A_n = \frac{2}{n\pi}\sin\left(\frac{n\pi\eta}{b}\right)\frac{\sinh[n\pi(a-\xi)/b]}{\sinh(n\pi a/b)}, \tag{15.6.60}$$

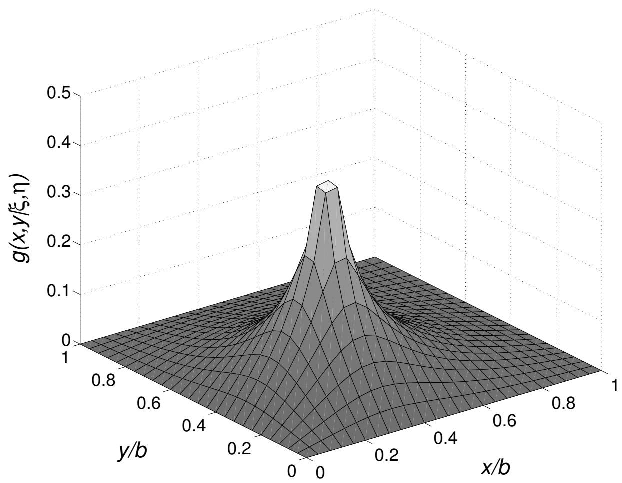


Figure 15.6.2: The Green's function, Equation 15.6.62 or Equation 15.6.63, for the planar Poisson equation over a rectangular area with Dirichlet boundary conditions on all sides when $a = b$ and $\xi/b = \eta/b = 0.3$.

and

$$B_n = \frac{2}{n\pi} \sin\left(\frac{n\pi\eta}{b}\right) \frac{\sinh(n\pi\xi/b)}{\sinh(n\pi a/b)}. \tag{15.6.61}$$

This yields the Green's function

$$g(x, y|\xi, \eta) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sinh[n\pi(a - x_>)/b] \sinh(n\pi x_</b)}{n \sinh(n\pi a/b)} \sin\left(\frac{n\pi\eta}{b}\right) \sin\left(\frac{n\pi y}{b}\right), \tag{15.6.62}$$

where $x_> = \max(x, \xi)$ and $x_< = \min(x, \xi)$. Figure 15.6.2 illustrates Equation 15.6.62 in the case of a square domain with $\xi/b = \eta/b = 0.3$.

If we began with a Fourier expansion in the y-direction, we would have obtained

$$g(x, y|\xi, \eta) = \frac{2}{\pi} \sum_{m=1}^{\infty} \frac{\sinh[m\pi(b - y_>)/a] \sinh(m\pi y_</a)}{m \sinh(m\pi b/a)} \sin\left(\frac{m\pi\xi}{a}\right) \sin\left(\frac{m\pi x}{a}\right), \tag{15.6.63}$$

where $y_> = \max(y, \eta)$ and $y_< = \min(y, \eta)$. □

• **Example 15.6.5: Two-dimensional Helmholtz equation over a rectangular domain**

The problem to be solved is

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} + k_0^2 g = -\delta(x - \xi)\delta(y - \eta), \tag{15.6.64}$$

where $0 < x, \xi < a$, and $0 < y, \eta < b$, subject to the boundary conditions that

$$g(0, y|\xi, \eta) = g(a, y|\xi, \eta) = g(x, 0|\xi, \eta) = g(x, b|\xi, \eta) = 0. \tag{15.6.65}$$

We use the same technique to solve Equation 15.6.64 as we did in the previous example by assuming that the Green's function has the form

$$g(x,y|\xi,\eta) = \sum_{m=1}^{\infty} G_m(x|\xi) \sin\left(\frac{m\pi y}{b}\right), \tag{15.6.66}$$

where the coefficients $G_m(x|\xi)$ are undetermined functions of x, ξ, and η. Substituting this series into Equation 15.6.64, multiplying by $2\sin(n\pi y/b)/b$, and integrating over y, we find that

$$\frac{d^2 G_n}{dx^2} - \left(\frac{n^2\pi^2}{b^2} - k_0^2\right) G_n = -\frac{2}{b} \sin\left(\frac{n\pi\eta}{b}\right) \delta(x-\xi). \tag{15.6.67}$$

The first method for solving Equation 15.6.67 involves writing

$$\delta(x-\xi) = \frac{2}{a} \sum_{m=1}^{\infty} \sin\left(\frac{m\pi\xi}{a}\right) \sin\left(\frac{m\pi x}{a}\right), \tag{15.6.68}$$

and

$$G_n(x|\xi) = \frac{2}{a} \sum_{m=1}^{\infty} a_{nm} \sin\left(\frac{m\pi x}{a}\right). \tag{15.6.69}$$

Upon substituting Equations 15.6.68 and 15.6.69 into Equation 15.6.67, we obtain

$$\begin{aligned} &\sum_{m=1}^{\infty} \left(k_0^2 - \frac{m^2\pi^2}{a^2} - \frac{n^2\pi^2}{b^2}\right) a_{nm} \sin\left(\frac{m\pi x}{a}\right) \\ &\qquad = -\frac{4}{ab} \sum_{m=1}^{\infty} \sin\left(\frac{n\pi\eta}{b}\right) \sin\left(\frac{m\pi\xi}{a}\right) \sin\left(\frac{m\pi x}{a}\right). \end{aligned} \tag{15.6.70}$$

Matching similar harmonics,

$$a_{nm} = \frac{4\sin(m\pi\xi/a)\sin(n\pi\eta/b)}{ab(m^2\pi^2/a^2 + n^2\pi^2/b^2 - k_0^2)}, \tag{15.6.71}$$

and the *bilinear form of the Green's function* is

$$g(x,y|\xi,\eta) = \frac{4}{ab} \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{\sin(m\pi\xi/a)\sin(n\pi\eta/b)\sin(m\pi x/a)\sin(n\pi y/b)}{m^2\pi^2/a^2 + n^2\pi^2/b^2 - k_0^2}. \tag{15.6.72}$$

See Figure 15.6.3. The bilinear form of the Green's function for the two-dimensional Helmholtz equation with Neumann boundary conditions is left as Problem 15.

As in the previous example, we can construct an alternative to the bilinear form of the Green's function, Equation 15.6.72, by writing Equation 15.6.67 as

$$\frac{d^2 G_n}{dx^2} - k_n^2 G_n = -\frac{2}{b} \sin\left(\frac{n\pi\eta}{b}\right) \delta(x-\xi), \tag{15.6.73}$$

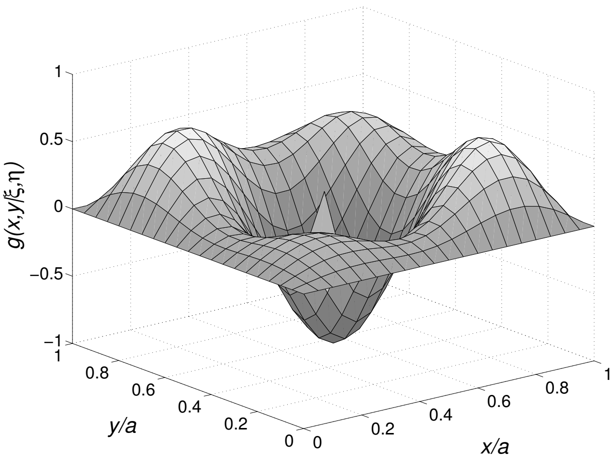


Figure 15.6.3: The Green's function, Equation 15.6.72, for Helmholtz's equation over a rectangular region with a Dirichlet boundary condition on the sides when $a = b$, $k_0 a = 10$, and $\xi/a = \eta/a = 0.35$.

where $k_n^2 = n^2\pi^2/b^2 - k_0^2$. The homogeneous solution to Equation 15.6.73 is now

$$G_n(x|\xi) = \begin{cases} A_n \sinh(k_n x), & 0 \le x \le \xi, \\ B_n \sinh[k_n(a-x)], & \xi \le x \le a. \end{cases} \tag{15.6.74}$$

This solution satisfies the boundary conditions at both endpoints.

Because $G_n(x|\xi)$ must be continuous at $x = \xi$,

$$A_n \sinh(k_n \xi) = B_n \sinh[k_n(a-\xi)]. \tag{15.6.75}$$

On the other hand, the jump discontinuity involving $G_n'(x|\xi)$ yields

$$-k_n B_n \cosh[k_n(a-\xi)] - k_n A_n \cosh(k_n \xi) = -\frac{2}{b}\sin\left(\frac{n\pi\eta}{b}\right). \tag{15.6.76}$$

Solving for A_n and B_n,

$$A_n = \frac{2}{bk_n}\sin\left(\frac{n\pi\eta}{b}\right)\frac{\sinh[k_n(a-\xi)]}{\sinh(k_n a)}, \tag{15.6.77}$$

and

$$B_n = \frac{2}{bk_n}\sin\left(\frac{n\pi\eta}{b}\right)\frac{\sinh(k_n\xi)}{\sinh(k_n a)}. \tag{15.6.78}$$

This yields the Green's function

$$\begin{aligned} g(x,y|\xi,\eta) = {} & \frac{2}{b}\sum_{n=1}^{N}\frac{\sin[\kappa_n(a-x_>)]\sin(\kappa_n x_<)}{\kappa_n \sin(\kappa_n a)}\sin\left(\frac{n\pi\eta}{b}\right)\sin\left(\frac{n\pi y}{b}\right) \\ & + \frac{2}{b}\sum_{n=N+1}^{\infty}\frac{\sinh[k_n(a-x_>)]\sinh(k_n x_<)}{k_n \sinh(k_n a)}\sin\left(\frac{n\pi\eta}{b}\right)\sin\left(\frac{n\pi y}{b}\right), \end{aligned} \tag{15.6.79}$$

where $x_> = \max(x, \xi)$ and $x_< = \min(x, \xi)$. Here N denotes the largest value of n such that $k_n^2 < 0$, and $\kappa_n^2 = k_0^2 - n^2\pi^2/b^2$. If we began with a Fourier expansion in the y direction, we would have obtained

$$g(x, y|\xi, \eta) = \frac{2}{a}\sum_{m=1}^{M}\frac{\sin[\kappa_m(b - y_>)]\sin(\kappa_m y_<)}{\kappa_m \sin(\kappa_m b)}\sin\left(\frac{m\pi\xi}{a}\right)\sin\left(\frac{m\pi x}{a}\right)$$
$$+\frac{2}{a}\sum_{m=M+1}^{\infty}\frac{\sinh[k_m(b - y_>)]\sinh(k_m y_<)}{k_m \sinh(k_m b)}\sin\left(\frac{m\pi\xi}{a}\right)\sin\left(\frac{m\pi x}{a}\right), \quad \textbf{(15.6.80)}$$

where M denotes the largest value of m such that $k_m^2 < 0$, $k_m^2 = m^2\pi^2/a^2 - k_0^2$, $\kappa_m^2 = k_0^2 - m^2\pi^2/a^2$, $y_< = \min(y, \eta)$, and $y_> = \max(y, \eta)$. □

• Example 15.6.6: Two-dimensional Helmholtz equation over a circular disk

In this example, we find the Green's function for the Helmholtz equation when the domain consists of the circular region $0 < r < a$. The Green's function is governed by the partial differential equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 g}{\partial\theta^2} + k_0^2 g = -\frac{\delta(r-\rho)\delta(\theta-\theta')}{r}, \quad \textbf{(15.6.81)}$$

where $0 < r, \rho < a$, and $0 \le \theta, \theta' \le 2\pi$, with the boundary conditions

$$\lim_{r\to 0}|g(r, \theta|\rho, \theta')| < \infty, \qquad g(a, \theta|\rho, \theta') = 0, \qquad 0 \le \theta, \theta' \le 2\pi. \quad \textbf{(15.6.82)}$$

The Green's function must be periodic in θ.

We begin by noting that

$$\delta(\theta - \theta') = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty}\cos[n(\theta-\theta')] = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\cos[n(\theta-\theta')]. \quad \textbf{(15.6.83)}$$

Therefore, the solution has the form

$$g(r, \theta|\rho, \theta') = \sum_{n=-\infty}^{\infty} g_n(r|\rho)\cos[n(\theta-\theta')]. \quad \textbf{(15.6.84)}$$

Substituting Equation 15.6.83 and Equation 15.6.84 into Equation 15.6.81 and simplifying, we find that

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dg_n}{dr}\right) - \frac{n^2}{r^2}g_n + k_0^2 g_n = -\frac{\delta(r-\rho)}{2\pi r}. \quad \textbf{(15.6.85)}$$

The solution to Equation 15.6.85 is the Fourier-Bessel series

$$g_n(r|\rho) = \sum_{m=1}^{\infty} A_{nm} J_n\left(\frac{k_{nm}r}{a}\right), \quad \textbf{(15.6.86)}$$

where k_{nm} is the mth root of $J_n(k) = 0$. Upon substituting Equation 15.6.86 into Equation 15.6.85 and solving for A_{nm}, we have that

$$(k_0^2 - k_{nm}^2/a^2)A_{nm} = -\frac{1}{\pi a^2 J'^2_n(k_{nm})}\int_0^a \delta(r-\rho)J_n\left(\frac{k_{nm}r}{a}\right)dr, \quad \textbf{(15.6.87)}$$

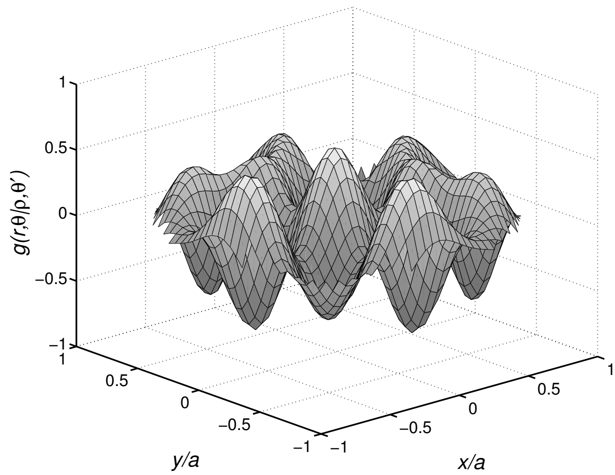


Figure 15.6.4: The Green's function, Equation 15.6.89, for Helmholtz's equation within a circular disk with a Dirichlet boundary condition on the rim when $k_0a = 10$, $\rho/a = 0.35\sqrt{2}$, and $\theta' = \pi/4$.

or

$$A_{nm} = \frac{J_n(k_{nm}\rho/a)}{\pi(k_{nm}^2 - k_0^2a^2)J'^2_n(k_{nm})}. \tag{15.6.88}$$

Thus, the Green's function[25] is

$$g(r,\theta|\rho,\theta') = \frac{1}{\pi}\sum_{n=-\infty}^{\infty}\sum_{m=1}^{\infty}\frac{J_n(k_{nm}\rho/a)J_n(k_{nm}r/a)}{(k_{nm}^2 - k_0^2a^2)J'^2_n(k_{nm})}\cos[n(\theta-\theta')]. \tag{15.6.89}$$

See Figure 15.6.4.

Problems

1. Using a Fourier sine expansion in the x-direction, construct the Green's function governed by the planar Poisson equation

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta), \quad 0 < x, \xi < a, \quad -\infty < y, \eta < \infty,$$

with the Dirichlet boundary conditions

$$g(0, y|\xi,\eta) = g(a, y|\xi,\eta) = 0, \quad -\infty < y < \infty,$$

[25] For an example of its use, see Zhang, D. R., and C. F. Foo, 1999: Fields analysis in a solid magnetic toroidal core with circular cross section based on Green's function. *IEEE Trans. Magnetics*, **35**, 3760–3762.

and the conditions at infinity that

$$\lim_{|y|\to\infty} g(x,y|\xi,\eta) \to 0, \qquad 0 < x < a.$$

2. Using a generalized Fourier expansion in the x-direction, construct the Green's function governed by the planar Poisson equation

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta), \quad 0 < x,\xi < a, \quad -\infty < y,\eta < \infty,$$

with the Neumann and Dirichlet boundary conditions

$$g_x(0,y|\xi,\eta) = g(a,y|\xi,\eta) = 0, \quad -\infty < y < \infty,$$

and the conditions at infinity that

$$\lim_{|y|\to\infty} g(x,y|\xi,\eta) \to 0, \qquad 0 < x < a.$$

3. Using a Fourier sine expansion in the y-direction, show that the Green's function governed by the planar Poisson equation

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta), \quad 0 < x,\xi < a, \quad 0 < y,\eta < b,$$

with the boundary conditions

$$g(x,0|\xi,\eta) = g(x,b|\xi,\eta) = 0,$$

and

$$g(0,y|\xi,\eta) = g_x(a,y|\xi,\eta) + \beta\, g(a,y|\xi,\eta) = 0, \qquad \beta \geq 0,$$

is

$$g(x,y|\xi,\eta) = \sum_{n=1}^{\infty} \frac{\sinh(\nu x_<)\left\{\nu\cosh\left[\nu(a-x_>)\right] + \beta\sinh\left[\nu(a-x_>)\right]\right\}}{\nu^2\cosh(\nu a) + \beta\nu\sinh(\nu a)} \sin\left(\frac{n\pi\eta}{b}\right)\sin\left(\frac{n\pi y}{b}\right),$$

where $\nu = n\pi/b$, $x_> = \max(x,\xi)$, and $x_< = \min(x,\xi)$.

4. Using the Fourier series representation of the delta function in a circular domain:

$$\delta(\theta-\theta') = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} \cos[n(\theta-\theta')], \qquad 0 \leq \theta,\theta' \leq 2\pi,$$

construct the Green's function governed by the planar Poisson equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 g}{\partial\theta^2} = -\frac{\delta(r-\rho)\delta(\theta-\theta')}{r},$$

where $a < r, \rho < b$, and $0 \le \theta, \theta' \le 2\pi$, subject to the boundary conditions $g(a, \theta|\rho, \theta') = g(b, \theta|\rho, \theta') = 0$ and periodicity in θ. You may want to review how to solve the equidimensional equation, Section 2.7.

5. Construct the Green's function governed by the planar Poisson equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} = -\frac{\delta(r-\rho)\delta(\theta-\theta')}{r},$$

where $0 < r, \rho < \infty$, and $0 < \theta, \theta' < \beta$, subject to the boundary conditions that $g(r, 0|\rho, \theta') = g(r, \beta|\rho, \theta') = 0$ for all r. Hint:

$$\delta(\theta - \theta') = \frac{2}{\beta}\sum_{n=1}^{\infty} \sin\left(\frac{n\pi\theta'}{\beta}\right) \sin\left(\frac{n\pi\theta}{\beta}\right).$$

You may want to review how to solve the equidimensional equation, Section 2.7.

6. Construct the Green's function governed by the planar Poisson equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial g}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} = -\frac{\delta(r-\rho)\delta(\theta-\theta')}{r},$$

where $0 < r, \rho < a$, and $0 < \theta, \theta' < \beta$, subject to the boundary conditions $g(r, 0|\rho, \theta') = g(r, \beta|\rho, \theta') = g(a, \theta|\rho, \theta') = 0$. Hint:

$$\delta(\theta - \theta') = \frac{2}{\beta}\sum_{n=1}^{\infty} \sin\left(\frac{n\pi\theta'}{\beta}\right) \sin\left(\frac{n\pi\theta}{\beta}\right).$$

You may want to review how to solve the equidimensional equation, Section 2.7.

7. Using a Fourier sine series in the z-direction and the results from Problem 31 at the end of Section 6.5, find the Green's function governed by the axisymmetric Poisson equation

$$\frac{\partial^2 g}{\partial r^2} + \frac{1}{r}\frac{\partial g}{\partial r} + \frac{\partial^2 g}{\partial z^2} = -\frac{\delta(r-\rho)\delta(z-\zeta)}{2\pi r},$$

where $0 < r, \rho < a$, and $0 < z, \zeta < L$, subject to the boundary conditions

$$g(r, 0|\rho, \zeta) = g(r, L|\rho, \zeta) = 0, \quad 0 < r < a,$$

and

$$\lim_{r \to 0} |g(r, z|\rho, \zeta)| < \infty, \quad g(a, z|\rho, \zeta) = 0, \quad 0 < z < L.$$

8. Following Example 15.6.5 except for using Fourier cosine series, construct the Green's function[26] governed by the planar Helmholtz equation

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} + k_0^2 g = -\delta(x-\xi)\delta(y-\eta), \quad 0 < x, \xi < a, \quad 0 < y, \eta < b,$$

[26] Kulkarni et al. (Kulkarni, S., F. G. Leppington, and E. G. Broadbent, 2001: Vibrations in several interconnected regions: A comparison of SEA, ray theory and numerical results. *Wave Motion*, **33**, 79–96) solved this problem when the domain has two different, constant k_0^2's.

subject to the Neumann boundary conditions

$$g_x(0,y|\xi,\eta) = g_x(a,y|\xi,\eta) = 0, \quad 0 < y < b,$$

and

$$g_y(x,0|\xi,\eta) = g_y(x,b|\xi,\eta) = 0, \quad 0 < x < a.$$

9. Using Fourier sine transforms,

$$g(x,y|\xi,\eta) = \frac{2}{\pi}\int_0^\infty G(k,y|\xi,\eta)\sin(kx)\,dk,$$

where

$$G(k,y|\xi,\eta) = \int_0^\infty g(x,y|\xi,\eta)\sin(kx)\,dx,$$

find the Green's function governed by

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta),$$

for the quarter space $0 < x, y$, with the boundary conditions

$$g(0,y|\xi,\eta) = g(x,0|\xi,\eta) = 0,$$

and

$$\lim_{x,y\to\infty} g(x,y|\xi,\eta) \to 0.$$

Step 1: Taking the Fourier sine transform in the x direction, show that the partial differential equation reduces to the ordinary differential equation

$$\frac{d^2G}{dy^2} - k^2 G = -\sin(k\xi)\delta(y-\eta),$$

with the boundary conditions

$$G(k,0|\xi,\eta) = 0, \qquad \text{and} \qquad \lim_{y\to\infty} G(k,y|\xi,\eta) \to 0.$$

Step 2: Show that the particular solution to the ordinary differential equation in Step 1 is

$$G_p(k,y|\xi,\eta) = \frac{\sin(k\xi)}{2k} e^{-k|y-\eta|}.$$

You may want to review Example 15.2.8.

Step 3: Find the homogeneous solution to the ordinary differential equation in Step 1 so that the general solution satisfies the boundary conditions. Show that the general solution is

$$G(k,y|\xi,\eta) = \frac{\sin(k\xi)}{2k}\left[e^{-k|y-\eta|} - e^{-k(y+\eta)}\right].$$

Step 4: Taking the inverse, show that

$$g(x,y|\xi,\eta) = \frac{1}{\pi}\int_0^\infty \left[e^{-k|y-\eta|} - e^{-k(y+\eta)}\right] \sin(k\xi)\,\sin(kx)\,\frac{dk}{k}.$$

Step 5: Performing the integration,[27] show that

$$g(x,y|\xi,\eta) = -\frac{1}{4\pi}\ln\left\{\frac{[(x-\xi)^2+(y-\eta)^2][(x+\xi)^2+(y+\eta)^2]}{[(x-\xi)^2+(y+\eta)^2][(x+\xi)^2+(y-\eta)^2]}\right\}.$$

10. Find the free-space Green's function[28] governed by

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} - g = -\delta(x-\xi)\delta(y-\eta), \qquad -\infty < x,y,\xi,\eta < \infty.$$

Step 1: Introducing the Fourier transform

$$g(x,y|\xi,\eta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(k,y|\xi,\eta)e^{ikx}\,dk,$$

where

$$G(k,y|\xi,\eta) = \int_{-\infty}^{\infty} g(x,y|\xi,\eta)e^{-ikx}\,dx,$$

show that the governing partial differential equation can be transformed into the ordinary differential equation

$$\frac{d^2G}{dy^2} - \left(k^2+1\right)G = -e^{-ik\xi}\delta(y-\eta).$$

Step 2: Introducing the Fourier transform in the y-direction,

$$G(k,y|\xi,\eta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{G}(k,\ell|\xi,\eta)e^{i\ell y}\,d\ell,$$

where

$$\overline{G}(k,\ell|\xi,\eta) = \int_{-\infty}^{\infty} G(k,y|\xi,\eta)e^{-i\ell y}\,dy,$$

solve the ordinary differential equation in Step 1 and show that

$$G(k,y|\xi,\eta) = \frac{e^{-ik\xi}}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\ell(y-\eta)}}{k^2+\ell^2+1}\,d\ell.$$

---

[27] Gradshteyn, I. S., and I. M. Ryzhik, 1965: *Table of Integrals, Series, and Products*. Academic Press, Section 3.947, Formula 1.

[28] For its use, see Geisler, J. E., 1970: Linear theory of the response of a two layer ocean to a moving hurricane. *Geophys. Fluid Dyn.*, **1**, 249–272.

Step 3: Complete the problem by showing that

$$g(x,y|\xi,\eta) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{ik(x-\xi)}e^{i\ell(y-\eta)}}{k^2+\ell^2+1}\,d\ell\,dk$$
$$= \frac{1}{4\pi^2}\int_{0}^{\infty}\int_{0}^{2\pi}\frac{e^{ir\kappa\cos(\theta-\varphi)}}{\kappa^2+1}\kappa\,d\theta\,d\kappa$$
$$= \frac{1}{2\pi}\int_{0}^{\infty}\frac{J_0(r\kappa)}{\kappa^2+1}\kappa\,d\kappa = \frac{K_0(r)}{2\pi},$$

where $r=\sqrt{(x-\xi)^2+(y-\eta)^2}$, $k=\kappa\cos(\theta)$, $\ell=\kappa\sin(\theta)$, $x-\xi=r\cos(\varphi)$, and $y-\eta=r\sin(\varphi)$. You may want to review the material around Equation 9.3.73 through Equation 9.3.75. You will need to use integral tables[29] to obtain the final result.

11. Find the free-space Green's function governed by

$$\frac{\partial^2 g}{\partial x^2}+\frac{\partial^2 g}{\partial y^2}-\frac{\partial g}{\partial x} = -\delta(x-\xi)\delta(y-\eta), \qquad -\infty < x,y,\xi,\eta < \infty.$$

Step 1: By introducing $\varphi(x,y|\xi,\eta)$ such that

$$g(x,y|\xi,\eta) = e^{x/2}\varphi(x,y|\xi,\eta),$$

show that the partial differential equation for $g(x,y|\xi,\eta)$ becomes

$$\frac{\partial^2 \varphi}{\partial x^2}+\frac{\partial^2 \varphi}{\partial y^2}-\frac{\varphi}{4} = -e^{-\xi/2}\delta(x-\xi)\delta(y-\eta).$$

Step 2: After taking the Fourier transform with respect to x of the partial differential equation in Step 1, show that it becomes the ordinary differential equation

$$\frac{d^2\Phi}{dy^2} - \left(k^2+\tfrac{1}{4}\right)\Phi = -e^{-\xi/2-ik\xi}\delta(y-\eta).$$

Step 3: Introducing the same transformation as in Step 3 of the previous problem, show that

$$\Phi(k,y|\xi,\eta) = \frac{e^{-\xi/2-ik\xi}}{2\pi}\int_{-\infty}^{\infty}\frac{e^{i\ell(y-\eta)}}{k^2+\ell^2+\frac{1}{4}}\,d\ell,$$

and

$$\varphi(x,y|\xi,\eta) = \frac{e^{-\xi/2}}{2\pi}K_0(\tfrac{1}{2}r),$$

where $r=\sqrt{(x-\xi)^2+(y-\eta)^2}$.

Step 4: Using the transformation introduced in Step 1, show that

$$g(x,y|\xi,\eta) = \frac{e^{(x-\xi)/2}}{2\pi}K_0(\tfrac{1}{2}r).$$

---

[29] Gradshteyn and Ryzhik, op. cit., Section 6.532, Formula 6.

15.7 GALERKIN METHOD

In the previous sections we developed various analytic expressions for Green's functions. We close this chapter by showing how to construct a numerical approximation.

In Sections 6.9 and 9.8 we showed how finite elements can be used to solve differential equations by introducing subdomains known as *finite elements* rather than a grid of nodal points. The solution is then represented within each element by an interpolating polynomial. Unlike finite difference schemes that are constructed from Taylor expansions, the theory behind finite elements introduces concepts from functional analysis and variational methods to formulate the algebraic equations.

There are several paths that lead to the same finite element formulation. The two most common techniques are the Galerkin and collocation methods. In this section we focus on the *Galerkin method*. This method employs a rational polynomial, called a *basis function*, that satisfies the boundary conditions.

We begin by considering the Sturm-Liouville problem governed by

$$\frac{d^2\psi_n}{dx^2} + \lambda_n\psi_n = 0, \qquad 0 < x < L, \tag{15.7.1}$$

subject to the boundary conditions $\psi_n(0) = \psi_n(L) = 0$. Although we could solve this problem exactly, we will pretend that we cannot. Rather, we will assume that we can express it by

$$\psi_n(x) = \sum_{j=1}^{N} \alpha_{nj} f_j(x), \tag{15.7.2}$$

where $f_j(x)$ is our *basis function*. Clearly, it is desirable that $f_j(0) = f_j(L) = 0$.

How do we compute α_{nj}? We begin by multiplying Equation 15.7.1 by $f_i(x)$ and integrating the resulting equation from 0 and L. This yields

$$\int_0^L f_i(x)\frac{d^2\psi_n}{dx^2}\,dx + \lambda_n \int_0^L f_i(x)\psi_n(x)\,dx = 0, \tag{15.7.3}$$

where $i = 1, 2, 3, \ldots, N$. Next, we substitute Equation 15.7.2 and find that

$$\sum_{j=1}^{N}\left[\int_0^L f_i(x)f_j''(x)\,dx + \lambda_n \int_0^L f_i(x)f_j(x)\,dx\right]\alpha_{nj} = 0. \tag{15.7.4}$$

We can write Equation 15.7.4 as

$$(A + \lambda_n B)\mathbf{d} = \mathbf{0}, \tag{15.7.5}$$

where

$$a_{ij} = \int_0^L f_i(x)f_j''(x)\,dx = -\int_0^L f_i'(x)f_j'(x)\,dx, \tag{15.7.6}$$

$$b_{ij} = \int_0^L f_i(x)f_j(x)\,dx, \tag{15.7.7}$$

and the vector $\mathbf{d}$ contains the elements α_{nj}.

There are several obvious choices for $f_j(x)$:

• **Example 15.7.1**

The simplest choice for $f_j(x) = \sin(j\pi x/L)$. If we select $N = 2$, Equation 15.7.2 becomes

$$\psi_n(x) = \alpha_{n1} \sin\left(\frac{\pi x}{L}\right) + \alpha_{n2} \sin\left(\frac{2\pi x}{L}\right). \qquad (15.7.8)$$

From Equation 15.7.6 and Equation 15.7.7,

$$a_{ij} = -\left(\frac{j\pi}{L}\right)^2 \int_0^L \sin\left(\frac{i\pi x}{L}\right) \sin\left(\frac{j\pi x}{L}\right)\, dx, \qquad i = 1, 2,\ j = 1, 2; \qquad (15.7.9)$$

and

$$b_{ij} = \int_0^L \sin\left(\frac{i\pi x}{L}\right) \sin\left(\frac{j\pi x}{L}\right)\, dx, \qquad i = 1, 2,\ j = 1, 2. \qquad (15.7.10)$$

Performing the integrations, $a_{12} = a_{21} = b_{12} = b_{21} = 0$, $a_{11} = -\pi^2/(2L)$, $a_{22} = -2\pi^2/L$, and $b_{11} = b_{22} = L/2$.

Returning to Equation 15.7.5, it becomes

$$\begin{pmatrix} -\pi^2/2 + \lambda_n L^2/2 & 0 \\ 0 & -2\pi^2 + \lambda_n L^2/2 \end{pmatrix} \begin{pmatrix} \alpha_{n1} \\ \alpha_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \qquad (15.7.11)$$

In order for Equation 15.7.11 to have a unique solution,

$$\begin{vmatrix} -\pi^2/2 + \lambda_n L^2/2 & 0 \\ 0 & -2\pi^2 + \lambda_n L^2/2 \end{vmatrix} = 0. \qquad (15.7.12)$$

Equation 15.7.12 yields $4\lambda_1 = \lambda_2 = 4\pi^2/L^2$.

In summary,

$$\psi_1(x) = \sin(\pi x/L), \qquad \lambda_1 = \pi^2/L^2; \qquad (15.7.13)$$

and

$$\psi_2(x) = \sin(2\pi x/L), \qquad \lambda_2 = 4\pi^2/L^2, \qquad (15.7.14)$$

with $\alpha_{12} = \alpha_{21} = 0$. Here we have chosen that $\alpha_{11} = \alpha_{22} = 1$. □

• **Example 15.7.2**

Another possible choice for $f_j(x)$ involves polynomials of the form $(1 - x/L)(x/L)^j$ with $j = 1, 2$. Unlike the previous example, we have *nonorthogonal* basis functions here. Note that $f_j(0) = f_j(L) = 0$. Therefore, Equation 15.7.2 becomes

$$\psi_n(x) = \alpha_{n1}(1 - x/L)(x/L) + \alpha_{n2}(1 - x/L)(x/L)^2. \qquad (15.7.15)$$

From Equation 15.7.6 and Equation 15.7.7,

$$a_{ij} = -\frac{1}{L^2} \int_0^L \left(1 - \frac{x}{L}\right) \left(\frac{x}{L}\right)^i \left[j(j-1) \left(\frac{x}{L}\right)^{j-2} - j(j+1) \left(\frac{x}{L}\right)^{j-1} \right] dx \qquad (15.7.16)$$

$$= \frac{1}{L} \left[\frac{j(j-1)}{i+j-1} - \frac{j(j-1)}{i+j} - \frac{j(j+1)}{i+j} + \frac{j(j+1)}{i+j+1} \right], \qquad (15.7.17)$$

with $i = 1, 2$ and $j = 1, 2$. Similarly,

$$b_{ij} = \int_0^L \left(1 - \frac{x}{L}\right)\left(\frac{x}{L}\right)^i \left(1 - \frac{x}{L}\right)\left(\frac{x}{L}\right)^j \, dx \tag{15.7.18}$$

$$= L\left[\frac{1}{i+j+1} - \frac{2}{i+j+2} + \frac{1}{i+j+3}\right]. \tag{15.7.19}$$

Performing the computations, $a_{11} = -1/(3L)$, $a_{12} = a_{21} = -1/(6L)$, $a_{22} = -2/(15L)$, $b_{11} = L/30$, $b_{12} = b_{21} = L/60$, and $b_{22} = L/105$.

Returning to Equation 15.7.5, it becomes

$$\begin{pmatrix} -1/3 + \lambda_n L^2/30 & -1/6 + \lambda_n L^2/60 \\ -1/6 + \lambda_n L^2/60 & -2/15 + \lambda_n L^2/105 \end{pmatrix} \begin{pmatrix} \alpha_{n1} \\ \alpha_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{15.7.20}$$

In order for Equation 15.7.20 to have a unique solution,

$$\begin{vmatrix} -1/3 + \lambda_n L^2/30 & -1/6 + \lambda_n L^2/60 \\ -1/6 + \lambda_n L^2/60 & -2/15 + \lambda_n L^2/105 \end{vmatrix} = 0. \tag{15.7.21}$$

Equation 15.7.21 yields $\lambda_1 L^2 = 10$ and $\lambda_2 L^2 = 42$. Note how close these values of λ are to those found in the previous example. Returning to Equation 15.7.20, we find that $\alpha_{11} = 1$, $\alpha_{12} = 0$, $\alpha_{22} = 1$, and $\alpha_{21} = -1/2$.

In summary,

$$\psi_1(x) = \left(1 - \frac{x}{L}\right)\frac{x}{L}, \quad \lambda_1 = \frac{10}{L^2}; \tag{15.7.22}$$

and

$$\psi_2(x) = -\frac{1}{2}\left(1 - \frac{x}{L}\right)\frac{x}{L} + \left(1 - \frac{x}{L}\right)\left(\frac{x}{L}\right)^2, \quad \lambda_2 = \frac{42}{L^2}. \tag{15.7.23}$$

Because $f_j(x)$ are linearly independent, their use in the Galerkin expansion is quite acceptable. However, because these functions are not particularly orthogonal, their usefulness will become more difficult as N increases. Consequently, the choice of orthogonal functions is often best. □

How do we employ the Galerkin technique to approximate Green's functions? We begin by considering the inhomogeneous heat conduction problem:

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = F(x,t), \qquad 0 < x < L, \quad 0 < t, \tag{15.7.24}$$

with the boundary conditions

$$u(0,t) = u(L,t) = 0, \qquad 0 < t, \tag{15.7.25}$$

and the initial condition $u(x,0) = 0$, $0 < x < L$.

Let us write the solution to this problem as

$$u(x,t) = \sum_{n=1}^{N} c_n(t)\psi_n(x)e^{-\lambda_n t}. \tag{15.7.26}$$

Direct substitution of Equation 15.7.26 into Equation 15.7.24, followed by multiplying the resulting equation by $f_i(x)$ and integrating from 0 to L, gives

$$\sum_{n=1}^{N} c_n(t)e^{-\lambda_n t}\int_0^L f_i(x)\frac{d^2\psi_n}{dx^2}\,dx - \sum_{n=1}^{N}\left(\frac{dc_n}{dt}-\lambda_n c_n\right)e^{-\lambda_n t}\int_0^L f_i(x)\psi_n(x)\,dx$$
$$= -\int_0^L f_i(x)F(x,t)\,dx. \tag{15.7.27}$$

Because

$$\frac{d^2\psi_n}{dx^2}+\lambda_n\psi_n=0, \tag{15.7.28}$$

Equation 15.7.27 simplifies to

$$\sum_{n=1}^{N}\frac{dc_n}{dt}e^{-\lambda_n t}\int_0^L f_i(x)\psi_n(x)\,dx = \int_0^L f_i(x)F(x,t)\,dx = F_i^*(t), \tag{15.7.29}$$

where $i=1,2,\ldots,N$.

We must now find c_n. We can write Equation 15.7.29 as

$$\sum_{n=1}^{N} e_{in}e^{-\lambda_n t}\frac{dc_n}{dt}\,dx = F_i^*(t), \tag{15.7.30}$$

where

$$e_{in}=\sum_{j=1}^{N}\alpha_{nj}b_{ji}. \tag{15.7.31}$$

Using linear algebra, we find that

$$e^{-\lambda_n t}\frac{dc_n}{dt}=\sum_{i=1}^{N}p_{ni}F_i^*(t), \tag{15.7.32}$$

where p_{ni} are the elements of an array $P=E^{-1}$ and $E=(DB)^T$. The arrays D and B consist of elements α_{ij} and b_{ij}, respectively. Solving Equation 15.7.32, we find that

$$c_n(t)=A_n+\sum_{i=1}^{N}p_{ni}\int_0^t F_i^*(\eta)e^{\lambda_n\eta}\,d\eta. \tag{15.7.33}$$

Because $u(x,0)=0$, $c_n(0)=0$ and $A_n=0$.

We are now ready to find our Green's function. Let us set $F(x,t)=\delta(x-\xi)\delta(t-\tau)$. Then $F_i^*(t)=f_i(\xi)\delta(t-\tau)$ and

$$c_n(t)=H(t-\tau)\sum_{i=1}^{N}p_{ni}f_i(\xi)e^{\lambda_n\tau}. \tag{15.7.34}$$

From Equation 15.7.2, Equation 15.7.26, and Equation 15.7.34, we obtain the final result that

$$g(x,t|\xi,\tau)=H(t-\tau)\sum_{n=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{N}\alpha_{nj}p_{ni}f_i(\xi)f_j(x)e^{-\lambda_n(t-\tau)}. \tag{15.7.35}$$

□

• **Example 15.7.3**

In Example 15.5.2, we solved the Green's function problem:

$$\frac{\partial g}{\partial t} - \frac{\partial^2 g}{\partial x^2} = \delta(x-\xi)\delta(t-\tau), \qquad (15.7.36)$$

with the boundary condition

$$g(0,t|\xi,\tau) = g(L,t|\xi,\tau) = 0, \qquad (15.7.37)$$

and the initial condition $g(x,0|\xi,\tau) = 0$. There we found the solution (Equation 15.5.34):

$$g(x,t|\xi,\tau) = H(t-\tau)\sum_{n=1}^{\infty} \psi_n(\xi)\psi_n(x)e^{-k_n^2(t-\tau)}, \qquad (15.7.38)$$

where we have the orthonormal eigenfunctions

$$\psi_n(x) = \sqrt{2/L}\,\sin(k_n x), \qquad k_n = n\pi/L. \qquad (15.7.39)$$

Let us use the basis function $f_j(x) = (1-x/L)(x/L)^j$ to find the approximate Green's function to Equation 15.7.36. Here $j = 1, 2, 3, \ldots, N$,

For $N = 2$, we showed in Example 15.7.2 that

$$B = L\begin{pmatrix} 1/30 & 1/60 \\ 1/60 & 1/105 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 \\ -1/2 & 1 \end{pmatrix}. \qquad (15.7.40)$$

Consequently,

$$BD = \frac{L}{840}\begin{pmatrix} 28 & 14 \\ 0 & 1 \end{pmatrix}, \qquad E = \frac{L}{840}\begin{pmatrix} 28 & 0 \\ 14 & 1 \end{pmatrix}. \qquad (15.7.41)$$

Using Gaussian elimination,

$$P = E^{-1} = \frac{1}{L}\begin{pmatrix} 30 & 0 \\ -420 & 840 \end{pmatrix}. \qquad (15.7.42)$$

Therefore, the two-term approximation to the Green's function, Equation 15.7.38, is

$$\begin{aligned} g(x,t|\xi,\tau) = {} & \frac{30}{L}\frac{x}{L}\left(1-\frac{x}{L}\right)\frac{\xi}{L}\left(1-\frac{\xi}{L}\right)\exp\left[-\frac{10(t-\tau)}{L^2}\right]H(t-\tau) \\ & + \left[\frac{210}{L}\frac{x}{L}\left(1-\frac{x}{L}\right)\frac{\xi}{L}\left(1-\frac{\xi}{L}\right) - \frac{420}{L}\frac{x}{L}\left(1-\frac{x}{L}\right)\left(\frac{\xi}{L}\right)^2\left(1-\frac{\xi}{L}\right)\right. \\ & \left. - \frac{420}{L}\left(\frac{x}{L}\right)^2\left(1-\frac{x}{L}\right)\frac{\xi}{L}\left(1-\frac{\xi}{L}\right) + \frac{840}{L}\left(\frac{x}{L}\right)^2\left(1-\frac{x}{L}\right)\left(\frac{\xi}{L}\right)^2\left(1-\frac{\xi}{L}\right)\right] \\ & \times \exp\left[-\frac{42(t-\tau)}{L^2}\right]H(t-\tau). \end{aligned} \qquad (15.7.43)$$

For $N > 2$, hand computations are very cumbersome and numerical computations are necessary. For a specific N, we first compute the arrays A and B via Equation 15.7.17 and Equation 15.7.19.

Table 15.7.1: The Value of $L^2\lambda_n$ for $n = 1, 2, \ldots, N$ as a Function of N.

n	Exact	$N=2$	$N=3$	$N=4$	$N=6$	$N=8$	$N=10$
1	9.87	10.00	9.87	9.87	9.87	9.87	9.87
2	39.48	42.00	42.00	39.50	39.48	39.48	39.48
3	88.83		102.13	102.13	89.17	88.83	88.83
4	157.91			200.50	159.99	157.96	157.91
5	246.74				350.96	254.42	247.04
6	355.31				570.53	376.47	356.65
7	483.61					878.88	531.55
8	631.65					1298.03	725.34
9	799.44						1850.98
10	986.96						2548.73

```
for j = 1:N
for i = 1:N
  A(i,j) = j*(j-1)/(i+j-1) - j*(j-1)/(i+j) ...
         + j*(j+1)/(i+j+1) - j*(j+1)/(i+j) ;
  B(i,j) = 1/(i+j+1) - 2/(i+j+2) + 1/(i+j+3);
end; end
```

Next we compute the λ_n's and corresponding eigenfunctions: `[v,d] = eig(A,-B)`. Table 15.7.1 gives $L^2\lambda_n$ for several values of N.

Once we found the eigenvalues and eigenvectors, we now compute the matrices D, E, and P. For convenience we have reordered the eigenvalues so that their numerical value increases with n. Furthermore, we have set α_{nn} equal to one for $n = 1, 2, \ldots, N$.

```
[lambda,ix] = sort(temp);
for i = 1:N
for j = 1:N
  D(i,j) = v(j,ix(i));
end; end
for i = 1:N
  denom = D(i,i);
  for j = 1:N
    D(i,j) = D(i,j) / denom;
end; end
E = transpose(D*B);
P = inv(E);
```

Having computed the matrices D and P, our final task is the computation of the Green's function using Equation 15.7.35. The MATLAB code is:

```
phi_i(1) = (1-xi)*xi;
for i = 2:N
  phi_i(i) = xi*phi_i(i-1);
end

for ii = 1:idim

  x = (ii-1)*dx;
  phi_j(1) = (1-x)*x;
```

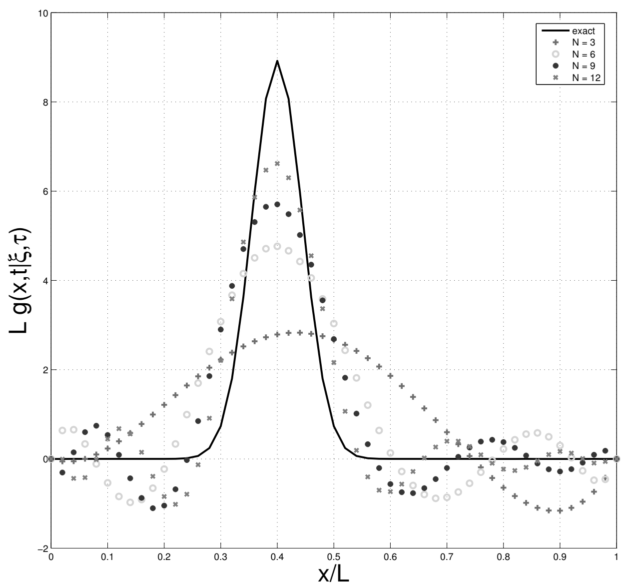


Figure 15.7.1: Comparison of the exact Green's function $Lg(x,t|\xi,\tau)$ for a one-dimensional heat equation given by Equation 15.7.38 (solid line) and approximate Green's functions found by the Galerkin method for different values of N. Here $(t-\tau)/L^2 = 0.001$ and $\xi = 0.4$.

```
  for j = 2:N
    phi_j(j) = x*phi_j(j-1);
  end

  for n = 1:N
  for j = 1:N
  for i = 1:N
    g(ii) = g(ii) + D(n,j).*P(n,i).*phi_j(j).*phi_i(i) ...
            .*exp(-lambda(n)*time);
end; end; end

end
```

In this code the parameter `time` denotes the quantity $(t-\tau)/L^2$. Figure 15.7.1 compares this approximate Green's function for various N against the exact solution. One of the problems with this method is finding the inverse of the array E. As N increases, the accuracy of the inverse becomes poorer.

Further Readings

Beck, J. V., K. D. Cole, A. Haji-Sheikh, and B. Litkouhi, 1992: *Heat Conduction Using Green's Functions.* Hemisphere Publishing Corp., 523 pp. Detailed study of solving heat conduction problems via Green's functions.

Carslaw, H. S., and J. C. Jaeger, 1959: *Conduction of Heat in Solids.* Oxford University Press, Chapter 14. An early classic for finding the Green's function for the heat equation.

Duffy, D. G., 2015: *Green's Functions with Applications.* Chapman & Hall/ CRC, 464 pp. A source book.

Özişik, M. N., 1993: *Heat Conduction.* John Wiley & Sons, Chapter 6. A book of how to solve partial differential equations of heat conduction.

Stakgold, I., 1997: *Green's Functions and Boundary Value Problems.* Wiley-Interscience, 720 pp. A systematic approach to boundary-value problems.

Chapter 16
Probability

So far in this book we presented mathematical techniques that are used to solve deterministic problems - problems in which the underlying physical processes are known exactly. In this and the next chapter we turn to problems in which uncertainty is key.

Although probability theory was first developed to explain the behavior of games of chance,[1] its usefulness in the physical sciences and engineering became apparent by the late nineteenth century. Consider, for example, the biological process of birth and death that we treated in Example 1.2.9. If b denotes the birth rate and d is the death rate, the size of the population $P(t)$ at time t is

$$P(t) = P(0)e^{(b-d)t}. \tag{16.0.1}$$

Let us examine the situation when $P(0) = 1$ and $b = 2d$ so that a birth is twice as likely to occur as a death. Then, Equation 16.0.1 predicts exponential growth with $P(t) = e^{dt}$. But the first event may be a death, a one-in-three chance since $d/(b + d) = 1/3$, and this would result in the population immediately becoming extinct. Consequently we see that for small populations, chance fluctuations become important and a deterministic model is inadequate.

The purpose of this and the next chapter is to introduce mathematical techniques that will lead to realistic models where chance plays an important role, and show under what conditions deterministic models will work. In this chapter we present those concepts that we will need in the next chapter to explain random processes.

---

[1] Todhunter, I., 1949: *A History of the Mathematical Theory of Probability From the Time of Pascal to That of Laplace*. Chelsea, 624 pp.; Hald, A., 1990: *A History of Probability and Statistics and Their Applications before 1750.* John Wiley & Sons, 586 pp.

16.1 REVIEW OF SET THEORY

Often we must count various objects in order to compute a probability. Sets provide a formal method to aid in these computations. Here we review important concepts from set theory.

Sets are collections of objects, such as the number of undergraduate students at a college. We define a set A either by naming the objects or describing the objects. For example, the set of natural numbers can be either enumerated:

$$A = \{1, 2, 3, 4, \ldots\}, \tag{16.1.1}$$

or described:

$$A = \{I : I \text{ is an integer and } I \geq 1\}. \tag{16.1.2}$$

Each object in set A is called an *element* and each element is *distinct*. Furthermore, the *ordering* of the elements within the set is not important.

Two sets are said to be equal if they contain the same elements and are written $A = B$. An element x of a set A is denoted by $x \in A$. A set with no elements is called a *empty* or *null* set and denoted by $\emptyset$. On the other hand, a *universal set* is the set of all elements under consideration.

A set B is *subset* of a set A, written $B \subset A$, if every element in B is also an element of A. For example, if $A = \{x : 0 \leq x < \infty\}$ and $S = \{x : -\infty < x < \infty\}$, then $A \subset S$. We can also use this concept to define the equality of sets A and B. Equality occurs when $A \subset B$ and $B \subset A$.

The *complement* of A, written $\overline{A}$, is the set of elements in S but not in A. For example, if $A = \{x : 0 \leq x < \infty\}$ and $S = \{x : -\infty < x < \infty\}$, then $\overline{A} = \{x : -\infty < x < 0\}$.

Two sets can be combined together to form a new set. This *union* of A and B, written $A \cup B$, creates a new set that contains elements that belong to A and/or B. This definition can be extended to multiple sets $A_1, A_2, \ldots, A_N$ so that the union is the set of elements for which each element belongs to at least one of these sets. It is written

$$A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_N = \bigcup_{i=1}^{N} A_i. \tag{16.1.3}$$

The *intersection* of sets A and B, written $A \cap B$, is defined as the set of elements that belong to both A and B. This definition can also be extended to multiple sets $A_1, A_2, \ldots, A_N$ so that the intersection is the set of elements for which each element belongs to all of these sets. It is written

$$A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_N = \bigcap_{i=1}^{N} A_i. \tag{16.1.4}$$

If two sets A and B have no elements in common, they are said to be *disjoint* and $A \cap B = \emptyset$.

A popular tool for visualizing set operations is the *Venn diagram*.[2] For sets A and B Figure 16.1.1 pictorially illustrates $A \cup B$, $A \cap B$, $\overline{A}$, and $A \cap \overline{B}$.

With these definitions a number of results follow: $\overline{\overline{A}} = A$, $A \cup \overline{A} = S$, $A \cap \overline{A} = \emptyset$, $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$, $A \cup S = S$, $A \cap S = A$, $\overline{S} = \emptyset$, and $\overline{\emptyset} = S$. Here S denotes the universal set.

Sets obey various rules similar to those encountered in algebra. They include:

---

[2] Venn, J., 2008: *Symbolic Logic.* Kessinger, 492 pp.

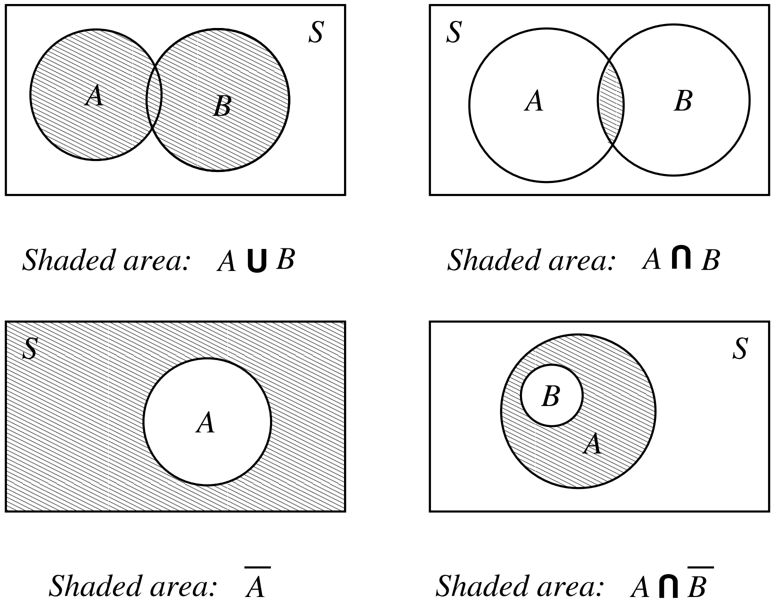


Figure 16.1.1: Examples of Venn diagrams for various configurations of sets A and B. Note that in the case of the lower right diagram, $B \subset A$.

1. Commutative properties: $A \cup B = B \cup A, \quad A \cap B = B \cap A$.
2. Associate properties: $A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C$.
3. Distributive properties: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
4. De Morgan's law: $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

Finally we define the *size* of a set. *Discrete* sets may have a finite number of elements or *countably infinite* number of elements. By countably infinite we mean that we could in theory count the number of elements in the sets. Two simple examples are $A = \{1, 2, 3, 4, 5, 6\}$ and $A = \{1, 4, 16, 64, \ldots\}$. Discrete sets lie in opposition to *continuous* sets where the elements are infinite in number and cannot be counted. A simple example is $A = \{x : 0 \leq x \leq 2\}$.

Problems

1. If $B \subset A$, use Venn diagrams to show that $A = B \cup (\overline{B} \cap A)$ and $B \cap (\overline{B} \cap A) = \emptyset$. Hint: Use the Venn diagram in the lower right frame of Figure 16.1.1.

2. Using Venn diagrams, show that $A \cup B = A \cup (\overline{A} \cap B)$ and $B = (A \cap B) \cup (\overline{A} \cap B)$. Hint: For $A \cap B$, use the upper right frame from Figure 16.1.1.

16.2 CLASSIC PROBABILITY

All questions of probability begin with the concept of an *experiment* where the governing principle is chance. The set of all possible outcomes of a random experiment is called the *sample space* (or universal set); we shall denote it by S. An element of S is called a *sample point*. The number of elements in S can be finite as in the flipping of a coin twice, infinite but countable such as repeatedly tossing a coin and counting the number of heads, or infinite and uncountable, as measuring the lifetime of a light bulb.

Any subset of the sample set S is called an *event*. If this event contains a single point, then the event is *elementary* or *simple*.

• Example 16.2.1

Consider an experiment that consists of two steps. In the first step, a die is tossed. If the number of dots on the top of the die is even, a coin is flipped; if the number of dots on the die is odd, a ball is selected from a box containing blue and green balls. The sample space is $S = \{1B, 1G, 2H, 2T, 3B, 3G, 4H, 4T, 5B, 5G, 6H, 6T\}$. The event A of obtaining a blue ball is $A = \{1B, 3B, 5B\}$, of obtaining a green ball is $B = \{1G, 3G, 5G\}$, and obtaining an even number of dots when the die is tossed is $C = \{2H, 2T, 4H, 4T, 6H, 6T\}$. The simple event of obtaining a 1 on the die and a blue ball is $D = \{1B\}$. □

Equally likely outcomes

An important class of probability problems consists of those whose outcomes are equally likely. This expression "equally likely" is essentially an intuitive one. For example, if a coin is tossed it seems reasonable that the coin is just as likely to fall "heads" as to fall "tails." Probability seeks to quantify this common sense.

Consider a sample space S of an experiment that consists of finitely many outcomes that are equally likely. Then the probability of an event A is

$$P(A) = \frac{\text{Number of points in } A}{\text{Number of points in } S}. \qquad (\mathbf{16.2.1})$$

With this simple definition we are ready to do some simple problems. An important aid in our counting is whether we can count a particular sample only once (sampling without replacement) or repeatedly (sampling with replacement). The following examples illustrate both cases.

• Example 16.2.2: Balls drawn from urns with replacement

Imagine the situation where we have an urn that has k red balls and $N - k$ black balls. A classic problem asks: What is the chance of two balls being drawn, one after another with replacement, where the first ball is red and the second one is black?

We begin by labeling the k red balls with $1, 2, 3, \ldots, k$ and black balls are numbered $k+1, k+2, \ldots, N$. The possible outcomes of the experiment can be written as a 2-tuple (z_1, z_2), where $z_1 \in 1, 2, 3, \ldots, N$ and $z_2 \in 1, 2, 3, \ldots, N$. A successful outcome is a red ball followed by a black one; we can express this case by $E = \{(z_1, z_2) : z_1 = 1, 2, \ldots, k; z_2 = k+1, k+2, \ldots, N\}$. Now the total number of 2-tuples in the sample space is N^2 while the total number of 2-tuples in E is $k(N-k)$. Therefore, the probability is

$$P(E) = \frac{k(N-k)}{N^2} = p(1-p), \qquad (\mathbf{16.2.2})$$

where $p = k/N$. □

• **Example 16.2.3: Balls drawn from urns without replacement**

Let us redo the previous example where the same ball now cannot be chosen twice. We can express this mathematically by the condition $z_1 \neq z_2$. The sample space has $N(N-1)$ balls and the number of successful 2-tuples is again $k(N-k)$. The probability is therefore given by

$$P(E) = \frac{k(N-k)}{N(N-1)} = \frac{k}{N}\frac{N-k}{N}\frac{N}{N-1} = p(1-p)\frac{N}{N-1}. \tag{16.2.3}$$

The restriction that we cannot replace the original ball has resulted in a higher probability. Why? We have reduced the number of red balls and thereby reduced the chance that we again selected another red ball while the situation with the black balls remains unchanged.

□

• **Example 16.2.4: The birthday problem**[3]

A classic problem in probability is: What is the chance that at least two individuals share the same birthday in a crowd of n people? Actually it is easier to solve the complementary problem: What is the chance that no one in a crowd of n individuals shares the same birthday?

For simplicity let us assume that there are only 365 days in the year. Each individual then has a birthday on one of these 365 days. Therefore, there are a total of $(365)^n$ possible outcomes in a given crowd.

Consider now each individual separately. The first person has a birthday on one of 365 days. The second person, who cannot have the same birthday, has one of the remaining 364 days. Therefore, if A denotes the event that no two people have the same birthday and each outcome is equally likely, then

$$P(A) = \frac{n(A)}{n(S)} = \frac{(365)(364)\cdots(365-n+1)}{(365)^n}. \tag{16.2.4}$$

To solve the original question, we note that $P(\overline{A}) = 1 - P(A)$ where $P(\overline{A})$ denotes the probability that at least two individuals share the same birthday.

If $n = 50$, $P(A) \approx 0.03$ and $P(\overline{A}) \approx 0.97$. On the other hand, if $n = 23$, $P(A) \approx 0.493$ and $P(\overline{A}) \approx 0.507$. Figure 16.2.1 illustrates $P(\overline{A})$ as a function of n. Nymann[4] computed the probability that in a group of n people, at least one pair will have the same birthday with at least one such pair among the first k people. □

In the previous examples we counted the objects in sets A and S. Sometimes we can define these sets only as areas on a graph. This graphical definition of probability is

$$P(A) = \frac{\text{Area covered by set } A}{\text{Area covered by set } S}. \tag{16.2.5}$$

The following example illustrates this definition.

---

[3] First posed by von Mises, R., 1939: Über Aufteilungs- und Besetzungswahrscheinlichkeiten. *Rev. Fac. Sci. Istambul*, **4**, 145–163.

[4] Nymann, J. E., 1975: Another generalization of the birthday problem. *Math. Mag.*, **53**, 111–125.

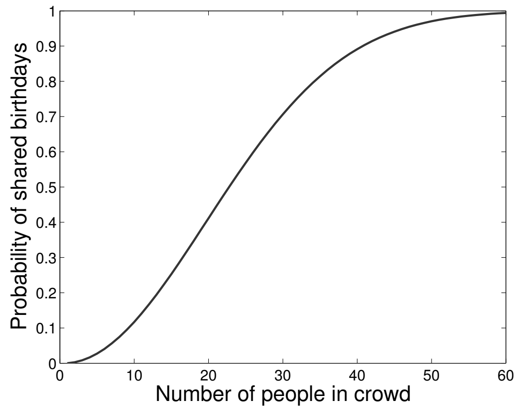


Figure 16.2.1: The probability that a pair of individuals in a crowd of n people share the same birthday.

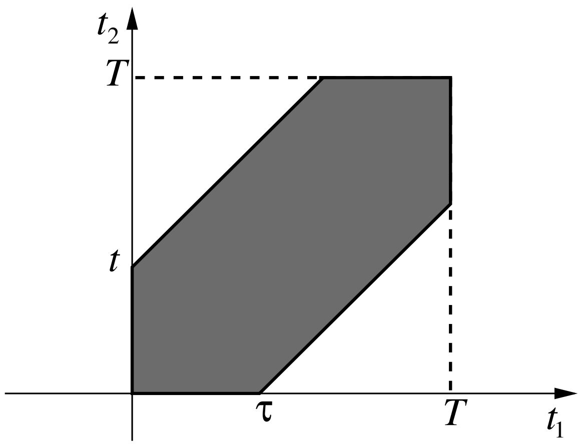


Figure 16.2.2: The graphical solution of whether two fellows can chat online between noon and T minutes after noon. The shaded area denotes the cases when the two will both be online whereas the rectangle gives the sample space.

• Example 16.2.5

Two friends, Joe and Dave, want to chat online but they will log on independently between noon and T minutes after noon. Because of their schedules Joe can only wait t minutes after his log-on while Dave can only spare τ minutes. Neither fellow can stay beyond T minutes after noon. What is the chance that they will chat?

Let us denote Joe's log-on time by t_1 and Dave's log-on time by t_2. Joe and Dave will chat if $0 < t_2 - t_1 < t$ and $0 < t_1 - t_2 < \tau$. In Figure 16.2.2 we show the situation where these inequalities are both satisfied. The area of the sample space is T^2. Therefore, from the geometrical definition of probability, the probability $P(A)$ that they will chat is

$$P(A) = \frac{T^2 - (T-t)^2/2 - (T-\tau)^2/2}{T^2}. \tag{16.2.6}$$

□

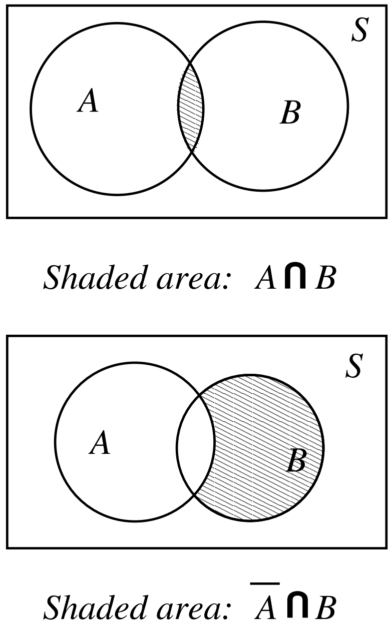


Figure 16.2.3: The Venn diagram used in the derivation of Property 5.

So far there has been a single event that interests us and we have only had to compute $P(A)$. Suppose we now have two events that we wish to follow. How are the probabilities $P(A)$ and $P(B)$ related?

Consider the case of flipping a coin. We could define event A as obtaining a head, $A = \{\text{head}\}$. Event B could be obtaining a tail, $B = \{\text{tail}\}$. Clearly $A \cup B = \{\text{head}, \text{tail}\} = S$, the sample space. Furthermore, $A \cap B = \emptyset$ and A and B are *mutually exclusive*. We already know that $P(A) = P(B) = \frac{1}{2}$. But what happens if $A \cap B$ is *not* an empty set?

From our definition of probability and previous examples, we introduce the following three basic axions:

Axion 1: $P(A), P(B) \geq 0$, Axion 2: $P(S) = 1$, Axion 3: $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$.

The first two axions are clearly true from the definition of probability and sample space. It is the third axion that needs some attention. Here we have two mutually exclusive events A and B in the sample space S. Because the number of points in $A \cup B$ equals the number of points in A plus the number of points in B, $n(A \cup B) = n(A) + n(B)$. Dividing both sides of this equation by the number of sample points and applying Equation 16.2.1, we obtain Axion 3 when $A \cap B = \emptyset$.

From these three axioms, the following properties can be written down:

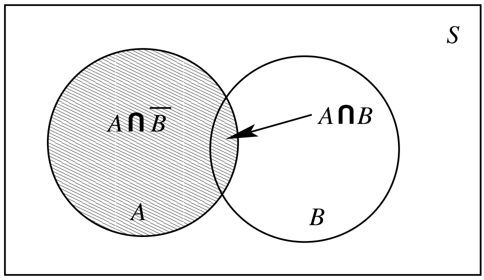


Figure 16.2.4: The Venn diagram that shows that $A = (A \cap \overline{B}) \cup (A \cap B)$.

1. $P(\overline{A}) = 1 - P(A)$
2. $P(\emptyset) = 0$
3. $P(A) \le P(B)$ if $A \subset B$
4. $P(A) \le 1$
5. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.

All of these properties follow readily from our definition of probability except for Property 5 and this is an important one. To prove this property from Axion 3, consider the Venn diagram shown in Figure 16.2.3. From this figure we see that

$$A \cup B = A \cup (\overline{A} \cap B) \qquad \text{and} \qquad B = (A \cap B) \cup (\overline{A} \cap B). \tag{16.2.7}$$

From Axion 3, we have that

$$P(A \cup B) = P(A) + P(\overline{A} \cap B), \tag{16.2.8}$$

and

$$P(B) = P(A \cap B) + P(\overline{A} \cap B). \tag{16.2.9}$$

Eliminating $P(\overline{A} \cap B)$ between Equation 16.2.8 and Equation 16.2.9, we obtain Property 5.

The following example illustrates a probability problem with two events A and B.

• Example 16.2.6

Consider Figure 16.2.4. From this figure, we see that $A = (A \cap \overline{B}) \cup (A \cap B)$. Because $A \cap \overline{B}$ and $A \cap B$ are mutually exclusive, then from Axion 3 we have that

$$P(A) = P(A \cap \overline{B}) + P(A \cap B). \tag{16.2.10}$$

□

• Example 16.2.7

A company has 400 employees. Every quarter, 100 of them are tested for drugs. The company's policy is to test everyone at random, whether they have been previously tested or not. What is the chance that someone is *not* tested?

The chance that someone *will* be tested is 1/4. Therefore, the chance that someone will *not* be tested is $1 - 1/4 = 3/4$. □

Permutations and combinations

By now it should be evident that your success at computing probabilities lies in correctly counting the objects in a given set. Here we examine two important concepts for systemic counting: permutations and combinations.

A *permutation* consists of ordering n objects *without any regard to their order*. For example, the six permutations of the three letters a, b, and c are abc, acb, bac, bca, cab, and cba. The number of permutations equals $n!$.

In a combination of given objects we select one or more objects without regard to their order. There are two types of combinations: (1) n different objects, taken k at a time, without repetition, and (2) n different objects, taken k at a time, with repetitions. In the first case, the number of sets that can be made up from n objects, each set containing k different objects and no two sets containing exactly the same k things, equals

$$\text{number of different combinations} = \binom{n}{k} \equiv \frac{n!}{k!(n-k)!}. \qquad \textbf{(16.2.11)}$$

Using the three letters a, b, and c, there are three combinations, taken two letters at a time, without repetition: ab, ac, and bc.

In the second case, the number of sets, consisting of k objects chosen from the n objects and each being used as often as desired, is

$$\text{number of different combinations} = \binom{n+k-1}{k}. \qquad \textbf{(16.2.12)}$$

Returning to our example using three letters, there are six combinations with repetitions: ab, ac, bc, aa, bb, and cc.

• Example 16.2.8

An urn contains r red balls and b blue balls. If a random sample of size m is chosen, what is the probability that it contains exactly k red balls?

If we choose a random sample of size m, we obtain $\binom{r+b}{m}$ possible outcomes. The number of samples that includes k red balls and $m-k$ blue balls is $\binom{r}{k}\binom{b}{m-k}$. Therefore, the probability that a sample of size m contains exactly k red balls is

$$\frac{\binom{r}{k}\binom{b}{m-k}}{\binom{r+b}{m}}. \qquad \square$$

• Example 16.2.9

A dog kennel has 50 dogs, including 5 German shepherds. (a) What is the probability of choosing 3 German shepherds if 10 dogs are randomly selected? (b) What is the probability of choosing all of the German shepherds in a group of 10 dogs that is chosen at random?

Let S denote the sample space of groups of 10 dogs. The number of those groups is $n(S) = 50!/(10!40!)$. Let A_i denote the set of 10 dogs that contain i German shepherds. Then the number of groups of 10 dogs that contain i German shepherds is $n(A_i) = 10!/[i!(10-i)!]$. Therefore, the probability that out of 50 dogs, we can select at random 10 dogs that include i German shepherds is

$$P(A_i) = \frac{n(A_i)}{n(S)} = \frac{10!10!40!}{i!(10-i)!50!}. \tag{16.2.13}$$

Thus, $P(A_3) = 1.1682 \times 10^{-8}$ and $P(A_5) = 2.453 \times 10^{-8}$. □

• Example 16.2.10

Consider an urn with n red balls and n blue balls inside. Let $R = \{r_1, r_2, \ldots, r_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$. Then the number of subsets of $R \cup B$ with n elements is $\binom{2n}{n}$. On the other hand, any subset of $R \cup B$ with n elements can be written as the union of a subset of R with i elements and a subset of B with $n-i$ elements for some $0 \le i \le n$. Because, for each i, there are $\binom{n}{i}\binom{n}{n-i}$ such subsets, the total number of subsets of red and blue balls with n elements equals $\sum_{i=0}^{n}\binom{n}{i}\binom{n}{n-i}$. Since both approaches must be equivalent,

$$\binom{2n}{n} = \sum_{i=0}^{n}\binom{n}{i}\binom{n}{n-i} = \sum_{i=0}^{n}\binom{n}{i}^2 \tag{16.2.14}$$

because $\binom{n}{n-i} = \binom{n}{i}$. □

Conditional probability

Often we are interested in the probability of an event A provided event B occurs. Denoting this *conditional probability* by $P(A|B)$, its probability is given by

$$P(A|B) = \frac{P(A\cap B)}{P(B)}, \qquad P(B) > 0, \tag{16.2.15}$$

where $P(A \cap B)$ is the joint probability of A and B. Similarly,

$$P(B|A) = \frac{P(A\cap B)}{P(A)}, \qquad P(A) > 0. \tag{16.2.16}$$

Therefore,

$$P(A\cap B) = P(A|B)P(B) = P(B|A)P(A), \tag{16.2.17}$$

and we obtain the famous *Bayes' rule*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{16.2.18}$$

• Example 16.2.11

Consider a box containing 10 pencils. Three of the pencils are defective with broken lead. If we draw 2 pencils out at random, what is the chance that we will have selected nondefective pencils?

There are two possible ways of selecting our two pencils: with and without replacement. Let Event A be that the first pencil is not defective and Event B be that the second pencil is not defective. Regardless of whether we replace the first pencil or not, $P(A) = \frac{7}{10}$ because each pencil is equally likely to be picked. If we then replace the first pencil, we have the same situation before any selection was made and $P(B|A) = P(A) = \frac{7}{10}$. Therefore,

$$P(A \cap B) = P(A)P(B|A) = 0.49. \qquad (\mathbf{16.2.19})$$

On the other hand, if we do not replace the first selected pencil, $P(B|A) = \frac{6}{9}$ because there is one fewer nondefective pencil. Consequently,

$$P(A \cap B) = P(A)P(B|A) = \frac{7}{10} \times \frac{6}{9} = \frac{14}{30} < 0.49. \qquad (\mathbf{16.2.20})$$

Why do we have a better chance of obtaining defective pencils if we don't replace the first one? Our removal of that first, nondefective pencil has reduced the uncertainty because we know that there are relatively more defective pencils in the remaining 9 pencils. This reduction in uncertainty must be reflected in a reduction in the chances that both selected pencils will be nondefective. □

Law of total probability

Conditional probabilities are useful because they allow us to simplify probability calculations. Suppose we have n mutually exclusive events $A_1, A_2, \ldots, A_n$ whose probabilities sum to unity, then

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n), \qquad (\mathbf{16.2.21})$$

where B is an arbitrary event, and $P(B|A_i)$ is the conditional probability of B assuming A_i. In other words, the law (or formula) of total probability expresses the total probability of an outcome that can be realized via several distinct events.

• Example 16.2.12

There are three boxes, each containing a different number of light bulbs. The first box has 10 bulbs, of which 4 are dead. The second has 6 bulbs, of which one is dead. Finally, there is a third box of eight bulbs, of which 3 bulbs are dead. What is the probability of choosing a dead bulb if a bulb is randomly chosen from one of the three boxes?

The probability of choosing a dead bulb is

$$P(D) = P(D|B_1)P(B_1) + P(D|B_2)P(B_2) + P(D|B_3)P(B_3) \qquad (\mathbf{16.2.22})$$

$$= \left(\frac{1}{3}\right)\left(\frac{4}{10}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{6}\right) + \left(\frac{1}{3}\right)\left(\frac{3}{8}\right) = \frac{113}{360}. \qquad (\mathbf{16.2.23})$$

If we had only one box with a total 24 bulbs, of which 8 were dead, then our chance of choosing a dead bulb would be $1/3 > 113/360$. □

Independent events

If events A and B satisfy the equation

$$P(A \cap B) = P(A)P(B), \tag{16.2.24}$$

they are called *independent events*. From Equation 16.2.15 and Equation 16.2.16, we see that if Equation 16.2.24 holds, then

$$P(A|B) = P(A), \qquad P(B|A) = P(B), \tag{16.2.25}$$

assuming that $P(A) \neq 0$ and $P(B) \neq 0$. Therefore, the term "independent" refers to the fact that the probability of A does *not* depend on the occurrence or non-occurrence of B, and vice versa.

• Example 16.2.13

Imagine some activity where you get two chances to be successful (for example, jumping for fruit still on a tree or shooting basketballs). If each attempt is independent and the probability of success 0.6 is the same for each trial, what is the probability of success after (at most) two tries?

There are two ways of achieving success. We can be successful in the first attempt with $P(S_1) = 0.6$ or we can fail and then be successful on the second attempt: $P(F_1 \cap S_2) = P(F_1)P(S_2) = (0.4)(0.6) = 0.24$, since each attempt is independent. Therefore, the probability of achieving success in two tries is $0.6 + 0.24 = 0.84$. Alternatively, we can compute the probability of failure in two attempts: $P(F_1 \cap F_2) = 0.16$. Then the probability of success with two tries would be the complement of the probability of two failures: $1 - 0.16 = 0.84$. □

• Example 16.2.14

Consider the tossing of a fair die. Let event A denote the tossing of a 2 or 3. Then $P(A) = P(\{2,3\}) = \frac{1}{3}$. Let event B denote tossing an odd number, $B = \{1,3,5\}$. Then $P(B) = \frac{1}{2}$.

Now $A \cap B = \{3\}$ and $P(A \cap B) = \frac{1}{6}$. Because $P(A \cap B) = P(A)P(B)$, events A and B are independent. □

Often we can characterize each outcome of an experiment consisting of n experiments as either a "success" or a "failure." If the probability of each individual success is p, then the probability of k successes and $n-k$ failures is $p^k(1-p)^{n-k}$. Because there are $n!/[k!(n-k)!]$ ways of achieving these k successes, the probability of an event having k successes in n independent trials is

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}, \tag{16.2.26}$$

where p is the probability of a success during one of the independent trials.

• Example 16.2.15

What is the probability of having two boys in a four-child family?

Let us assume that the probability of having a male is 0.5. Taking the birth of one child as a single trial,

$$P_4(2) = \frac{4!}{2!2!}\left(\frac{1}{2}\right)^4 = \frac{3}{8}. \tag{16.2.27}$$

Note that this is *not* 0.5, as one might initially guess.

Problems

1. For the following experiments, describe the sample space:

(a) flipping a coin twice.

(b) selecting two items out of three items $\{a, b, c\}$ without replacement.

(c) selecting two items out of three items $\{a, b, c\}$ with replacement.

(d) selecting three balls, one by one, from a box that contains four blue balls and five green balls without replacement.

(e) selecting three balls, one by one, from a box that contains four blue balls and five green balls with replacement.

2. Consider two fair dice. What is the probability of throwing them so that the dots sum to seven?

3. In throwing a fair die, what is the probability of obtaining a one *or* two on the top side of the cube?

4. What is the probability of getting heads exactly (a) twice or (b) thrice if you flip a fair coin 6 times?

5. An urn contains six red balls, three blue balls, and two green balls. Two balls are randomly selected. What is the sample space for this experiment? Let X denote the number of green balls selected. What are the possible values of X? Calculate $P(X = 1)$.

6. Consider an urn with 30 blue balls and 50 red balls in it. These balls are identical except for their color. If they are well mixed and you draw 3 balls without replacement, what is the probability that the balls are all of the same color?

7. A deck of cards has 52 cards, including 4 jacks and 4 ten's. What is the probability of selecting a jack *or* ten?

8. Two boys and two girls take their place on a stage to receive an award. What is the probability that the boys take the two end seats?

9. A lottery consists of posting a 3-digit number given by selecting 3 balls from 10 balls, each ball having the number from 1 to 10. The balls are not replaced after they are drawn. What are your chances of winning the lottery if the order does not matter? What are your chances of winning the lottery if the order does matter? Write a short MATLAB code and verify your results. You may want to read about the MATLAB intrinsic function `randperm`.

10. A circle of radius 1 is inscribed in a square with sides of length 2. A point is selected at random in the square in such a manner that all the subsets of equal area of the square are equally likely to contain the point. What is the probability that it is inside the circle?

11. In a rural high school, 20% of the students play football and 10% of them play football and wrestle. If Ed, a randomly selected student of this high school, played football, what is the probability that he also wrestles for his high school?

12. You have a well-shuffled card deck. What is the probability the second card in the deck is an ace?

13. We have two urns: One has 4 red balls and 6 green balls, the other has 6 red and 4 green. We toss a fair coin. If heads, we pick a random ball from the first urn, if tails from the second. What is the probability of getting a red ball? How do your results compare with the probability of getting a red ball if all of the red and green balls had been placed into a single urn?

14. A customer decides between two dinners: a "cheap" one and an "expensive" one. The probability that the customer chooses the expensive meal is $P(E) = 0.2$. A customer who chooses the expensive meal likes it with a 80% probability $P(L|E) = 0.8$. A customer who chooses the cheap meal dislikes it with 70% probability $P(D|C) = 0.7$.

(a) Compute the probability that a customer (1) will choose a cheap meal, (2) will be disappointed with an expensive meal, and (3) will like the the cheap meal.

(b) Use the law of total probability to compute the probability that a customer will be disappointed.

(c) If a customer found his dinner to his liking, what is the probability that he or she chose the expensive meal? Hint: Use Bayes' theorem.

15. Suppose that two points are *randomly* and *independently* selected from the interval $(0, 1)$. What is the probability the first one is greater than $1/4$, and the second one is less than $3/4$? Check your result using `rand` in MATLAB.

16. A certain brand of electronics chip is found to fail prematurely in 1% of all cases. If three of these chips are used in three independent sets of equipment, what is the probability that (a) all three will fail prematurely, (b) that two will fail prematurely, (c) that one will fail prematurely, and (d) that none will fail?

Project: Experimenting with MATLAB's Intrinsic Function `rand`

The MATLAB function `rand` can be used in simulations where sampling occurs with replacement. If we write `X = rand(1,100)`, the vector `X` contains 100 elements whose values vary between 0 and 1. Therefore, if you wish to simulate a fair die, then we can set up the following table:

```
 0 < X < 1/6      die with one dot showing
1/6 < X < 1/3     die with two dots showing
1/3 < X < 1/2     die with three dots showing
1/2 < X < 2/3     die with four dots showing
2/3 < X < 5/6     die with five dots showing
 5/6 < X < 1      die with six dots showing.
```

We can then write MATLAB code that counts the number of times that we obtain a one or two. Call this number `n`. Then the probability that we would obtain one or two dots on a fair die is `n/100`. Carry out this experiment and compare your answer with the result from Problem 2. What occurs as you do more and more experiments?

Project: Experimenting with
MATLAB's Intrinsic Function `randperm`

MATLAB's intrinsic function `randperm(m)` creates a random ordering of the numbers from `1` to `m`. If you execute `perm = randperm(365)`, this would produce a vector of length 365 and each element has a value lying between 1 and 365. If you repeat the process, you would obtain another list of 365 numbers but they would be in a different order.

Let us simulate the birthday problem. Invoking the `randperm` command, use the first element to simulate the birthday of student 1 in a class of `N` students. Repeatedly invoking this command, create vector `birthdays` that contains the birthdays of the `N` students. Then find out if any of the days are duplicates of another. (Hint: You might want to explore the MATLAB command `unique`.) Repeating this experiment many times, compute the chance that a class of size `N` has at least two students that have the same birthday. Compare your results with Equation 16.2.4. What occurs as the number of experiments increases?

16.3 DISCRETE RANDOM VARIABLES

In the previous section we presented the basic concepts of probability. In high school algebra you were introduced to the concept of a variable - a quantity that could vary unlike constants and parameters. Here we extend this idea to situations where the variations are due to randomness.

A *random variable* is a single-valued real function that assigns a real number, the *value*, to each sample point t of S. The variable can be discrete, such as the flipping of a coin, or continuous, such as the lifetime of a light bulb. The sample space S is the *domain* of the *random variable* $X(t)$, and the collection of all numbers $X(t)$ is the *range*. Two or more sample points can give the same value of $X(t)$ but we will never allow two different numbers in the range of $X(t)$ for a given t.

The term "random variable" is probably a poor one. Consider the simple example of tossing a coin. A random variable that describes this experiment is

$$X[s_i] = \begin{cases} 1, & s_1 = \text{head}, \\ 0, & s_2 = \text{tail}. \end{cases} \qquad \textbf{(16.3.1)}$$

An obvious question is: What is random about Equation 16.3.1? If a head is tossed, we obtain the answer one; if a tail is tossed, we obtain a zero. Everything is well defined; there is no element of chance here. The randomness arises from the tossing of the coin. Until the experiment (tossing of the coin) is performed, we do not know the outcome of the experiment and the value of the random variable. Therefore, *a random variable is a*

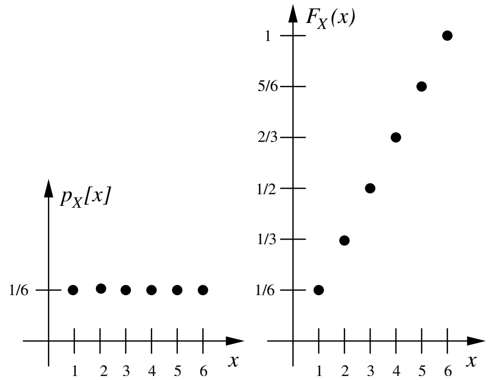


Figure 16.3.1: The probability mass function for a fair die.

variable that may take different values if a random experiment is conducted and its value is not known in advance.

We begin our study of random variables by focusing on those arising from discrete events. If X is discrete, X assumes only finitely many or countably many values: $x_1, x_2, x_3, \ldots$. For each possible value of x_i, there is a corresponding positive probability $p_X[x_1] = P(X = x_1), p_X[x_2] = P(X = x_2), \ldots$ given by the *probability mass function*. For values of x different from x_i, say $x_1 < x < x_2$, the probability mass function equals zero. Therefore, we have that

$$p_X[x_i] = \begin{cases} p_i, & x = x_i, \\ 0, & \text{otherwise}, \end{cases} \tag{16.3.2}$$

where $i = 1, 2, 3, \ldots$. A family of discrete random variables having the same probability mass family is called *identically distributed*. ‘

• Example 16.3.1

Consider a fair die. We can describe the results from rolling this fair die via the discrete random variable X, which has the possible values $x_i = 1, 2, 3, 4, 5, 6$ with the probability $p_X[x_i] = \frac{1}{6}$ each. Note that $0 \le p_X[x_i] < 1$ here. Furthermore,

$$\sum_{i=1}^{6} p_X[x_i] = 1. \tag{16.3.3}$$

Figure 16.3.1 illustrates the probability mass function. □

• Example 16.3.2

Let us now modify Example 16.3.1 so that

$$X[s_i] = \begin{cases} 1, & s_i = 1, 2, \\ 2, & s_i = 3, 4, \\ 3, & s_i = 5, 6. \end{cases} \tag{16.3.4}$$

The probability mass function becomes

$$p_X[1] = p_X[2] = p_X[3] = \tfrac{1}{3}. \tag{16.3.5}$$

□

Some Properties of the Probability Mass Function $p_X[x_i]$

$$0 \le p_X[x_k] < 1, \qquad p_X[x] = 0 \quad \text{if} \quad x \neq x_k, \qquad k = 1, 2, \ldots$$

$$\sum_n p_X[x_n] = 1$$

$$F_X(x) = P(X \le x) = \sum_{x_k \le x} p_X[x_k]$$

$$P(a < x \le b) = \sum_{a < x_k \le b} p_X[x_k]$$

• **Example 16.3.3**

Consider the probability mass function:

$$p_X[x_n] = \begin{cases} k(1/2)^n, & n = 0, 1, 2, \ldots, \\ 0, & \text{otherwise.} \end{cases} \tag{16.3.6}$$

Let us (a) find the value of k, (b) find $P(X = 2)$, (c) find $P(X \le 2)$, and (d) $P(X \ge 1)$.

From the properties of probability mass function,

$$k \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = k \frac{1}{1 - \frac{1}{2}} = 2k = 1. \tag{16.3.7}$$

Therefore, $k = \frac{1}{2}$. Note that $0 \le p_X[x_n] \le 1$.

Having found k, we immediately have

$$P(X = 2) = p_X[x_2] = \tfrac{1}{8}, \tag{16.3.8}$$

$$P(X \le 2) = p_X[x_0] + p_X[x_1] + p_X[x_2] = \tfrac{7}{8}, \tag{16.3.9}$$

and

$$P(X \ge 1) = 1 - P(X = 0) = \tfrac{1}{2}. \tag{16.3.10}$$

□

Having introduced the probability mass function, an alternative means of describing the probabilities of a discrete random variable is the *cumulative distribution function.* It is defined as

$$F_X(x) = P(X \le x), \qquad -\infty < x < \infty. \tag{16.3.11}$$

It is computed via

$$F_X(x) = \sum_{x_i \le x} p_X[x_i] = \sum_{x_i \le x} p_i. \tag{16.3.12}$$

Consequently, combining Equation 16.3.11 and Equation 16.3.12, we obtain

$$P(a < x \le b) = \sum_{a < x_i \le b} p_i. \tag{16.3.13}$$

Equation 16.3.13 gives the probability over the interval $(a, b]$.

• Example 16.3.4

A Bernoulli experiment is a random experiment, the outcome of which is a success or failure. Consider now a sequence of independent Bernoulli trials with probability p of success from trial to trial. This sequence is observed until the first success occurs. Let X denote a random variable that equals the trial number on which the first success occurs. The probability mass function is then

$$p_X[x_n] = (1-p)^{n-1}p, \qquad n = 1, 2, 3, \ldots. \tag{16.3.14}$$

Let us compute the cumulative distribution function.

For geometric series, we begin by noting that

$$\sum_{n=0}^{\infty} ar^n = \sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r}, \qquad |r| < 1. \tag{16.3.15}$$

Next we check Equation 16.3.14 and determine whether it is a valid probability mass function. It is because

$$\sum_{n=1}^{\infty} p_X[x_n] = \sum_{n=1}^{\infty} (1-p)^{n-1}p = \frac{p}{1-(1-p)} = 1, \tag{16.3.16}$$

where we used Equation 16.3.15. Next, we note that

$$P(X > m) = \sum_{n=m+1}^{\infty} (1-p)^{n-1}p = \frac{(1-p)^m p}{1-(1-p)} = (1-p)^m. \tag{16.3.17}$$

Therefore,

$$F_X(x) = P(X \le m) = 1 - P(X > m) = 1 - (1-p)^m, \tag{16.3.18}$$

where $m = 1, 2, 3, \ldots$. □

• Example 16.3.5: Generating discrete random variables via MATLAB

In this example we show how to generate a discrete random variable using MATLAB's intrinsic function `rand`. This MATLAB command produces random, uniformly distributed (equally probable) reals over the interval $(0, 1)$. How can we use this function when in the case of discrete random variables we have only integer values, such as $k = 1, 2, 3, 4, 5, 6$, in the case of tossing a die?[5]

Consider the Bernoulli random variable $X = k$, $k = 0, 1$. As you will show in your homework, it has the cumulative distribution function of

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1-p, & 0 \le x < 1, \\ 1, & 1 \le x. \end{cases} \tag{16.3.19}$$

[5] This technique is known as the inverse transform sampling method. See pages 85–102 in Devroye, L., 1986: *Non-Uniform Random Variable Generation*. Springer-Verlag, 843 pp.

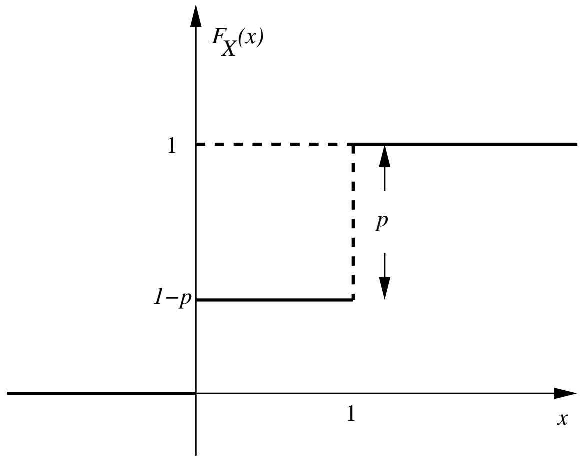


Figure 16.3.2: The cumulative distribution function for a Bernoulli random variable.

See Figure 16.3.2.

Imagine now a program that includes the MATLAB function `rand`, which yields the value `t`. Then, if $0 < t \leq 1 - p$, Figure 16.3.2 gives us that $X = 0$. On the other hand, if $1 - p < t < 1$, then $X = 1$. Thus, to obtain `M` realizations of the Bernoulli random variable `X`, the MATLAB code would read for a given `p`:

```
clear;
for i = 1:M
  t = rand(1,1);
    if (t <= 1-p) X(i,1) = 0;
    else
    X(i,1) = 1;
end; end
```

The end product of this code creates a vector `X` of length `M` consisting of a random variable with either zeros or ones. This is shown in Figure 16.3.3(a) when $p = 0.4$.

Once we have generated this random variable, we can use its relative frequency to compute its probability mass function and cumulative distribution function from

$$\hat{p}_X[x_k] = \frac{\text{Number of outcomes equal to } k}{M}, \tag{16.3.20}$$

and

$$\hat{F}_X(x) = \frac{\text{Number of outcomes } \leq x}{M}. \tag{16.3.21}$$

In Figure 16.3.3(b) we have computed the value of $\hat{p}_X[1]$. Clearly it should equal p. As this figure shows, we obtain poor results when M is small, with $\hat{p}_X[1]$ moving randomly above and below the correct answer. As M becomes larger, our estimate improves.

Problems

1. The Bernoulli distribution has the probability mass function

$$p_X[x_k] = P(X = k) = p^k(1-p)^{1-k}, \qquad k = 0, 1,$$

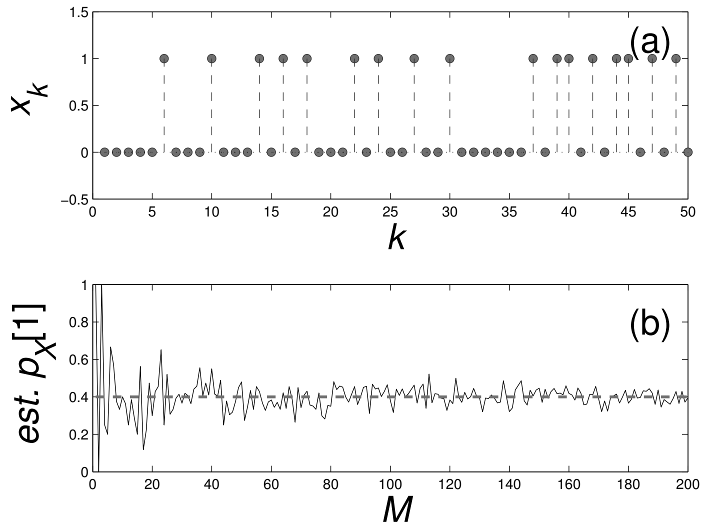


Figure 16.3.3: (a) Outcomes of the Bernoulli random variable generated by the MATLAB function `rand`. (b) The computed value of the probability mass function $p_X[1]$ as a function of M realization of the Bernoulli random variable. The dashed line is the line for the exact answer $p = 0.4$.

where $0 \le p \le 1$. (a) Show that this distribution is a valid probability mass function. (b) Find its cumulative distribution function.

2. An experiment is performed where a digit, ranging from 0 to 9, is repeatedly and randomly chosen. If X denotes the times that this experiment must be repeated until the digit 0 is selected, find $P(X)$.

3. A scientific company needs a programmer who knows an unusual programming language. If only 5% of programmers know this language, how many programmers should the company interview to have a 75% chance of finding such a programmer?

16.4 CONTINUOUS RANDOM VARIABLES

In the previous section we examined random variables that can assume only certain discrete values. Here we extend the concept of random variables so that they can take on values over a continuous interval. Typical examples of continuous random variables include the noisy portion of the voltage within an amplifier, the phase of a propagating wave, and the amount of precipitation.

An important quantity that we introduced in the previous section was the probability mass function. What is the corresponding function for continuous random variables? From the fundamental concepts of probability, we know that the probability of a continuous variable assuming one specific value out of its possible range values equals zero; it is merely one point out of an infinite number of points in the sample space. On the other hand, there is a finite probability that the value assumed by the random variable X will lie within an arbitrarily small interval dx and this probability will depend on the length of the interval.

Another factor that should influence the probability is the value of x. There is no reason why the probability of X should be independent of x. Consequently, an equation for probability in the interval $x < X \le x + dx$ requires a function $p_X(x)$, which acts as a

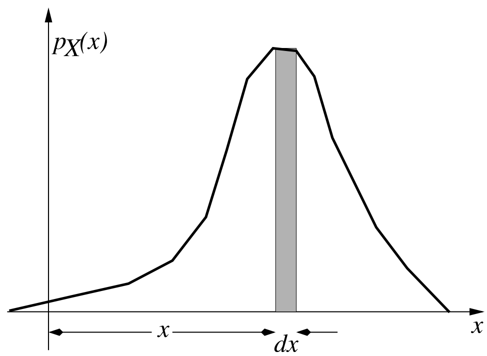


Figure 16.4.1: A probability density function.

weighting function and models the relative frequency behavior of X. For these reasons, the probability that a continuous random variable X will assume a value lying between x and $x + dx$ is given by

$$P(x < X \leq x + dx) = p_X(x)\,dx. \qquad \textbf{(16.4.1)}$$

Figure 16.4.1 illustrates a possible example of $p_X(x)$ where the shaded area equals the probability $P(x < X \leq x + dx)$. Clearly the function $p_X(x) = P(x < X \leq x + dx)/dx$ has the dimension of probability per infinitesimal interval dx and is called, for that reason, the *probability density*. Furthermore, although $p_X(x)\,dx \leq 1$, this does *not* mean that $p_X(x) \leq 1$. A family of random variables having the same probability density is *identically distributed*.

The function $p_X(x)$ must also satisfy several additional conditions. Because probability cannot be negative, $p_X(x) \geq 0$ of all x. Furthermore, as Figure 16.4.1 suggests, if we add up all of the possible values of x, then we have a certain event. We can express this mathematically by

$$\int_{-\infty}^{\infty} p_X(x)\,dx = 1. \qquad \textbf{(16.4.2)}$$

Thus, a probability density has the properties given by Equation 16.4.1 and Equation 16.4.2. It must also be a single-valued function of x. Note that these conditions do not require that $p_X(x)$ is a continuous function of x.

Let us now consider the probability $P(a < X \leq b)$ where a and b are constants. If we subdivide the range of x between a and b into infinitesimal intervals $(x, x + dx)$, the probability that the random variable will assume a value from one such interval is given by Equation 16.4.1. The probability that the variable will assume a value in the interval (a, b) equals the sum of the probabilities from each subinterval between a and b and is given by the area under the curve $p(x)$ between $x = a$ and $x = b$. Therefore,

$$P(a < X \leq b) = \int_a^b p_X(x)\,dx. \qquad \textbf{(16.4.3)}$$

If $a = -\infty$, we have that

$$P(X \leq b) = \int_{-\infty}^b p_X(x)\,dx. \qquad \textbf{(16.4.4)}$$

Alternatively, setting $b = \infty$,

$$P(a < X) = \int_a^{\infty} p_X(x)\,dx. \qquad \textbf{(16.4.5)}$$

Some Properties of the Probability Density Function $p_X(x)$

$$p_X(x) \geq 0, \qquad \int_{-\infty}^{\infty} p_X(x)\,dx = 1$$

$$P(a < X \leq b) = \int_a^b p_X(x)\,dx$$

From Equation 16.4.3 we also have

$$P(X > a) = 1 - P(X < a) = 1 - \int_{-\infty}^{a} p_X(x)\,dx = \int_a^{\infty} p_X(x)\,dx. \tag{16.4.6}$$

From Equation 16.4.4 we now define

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} p_X(\xi)\,d\xi. \tag{16.4.7}$$

This function $F_X(x)$ is called the *cumulative distribution function*, or simply the distribution function, of the random variable X. Clearly,

$$p_X(x) = F'_X(x). \tag{16.4.8}$$

Therefore, from the properties of $p_X(x)$, we have that (1) $F_X(x)$ is a nondecreasing function of x, (2) $F_X(-\infty) = 0$, (3) $F_X(\infty) = 1$, and (4) $P(a < X \leq b) = F_X(b) - F_X(a)$.

• Example 16.4.1

The continuous random variable X has the probability density function

$$p_X(x) = \begin{cases} k\left(x - x^2\right), & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{16.4.9}$$

What must be the value of k? What is the cumulative distribution function? What is $P(X < 1/2)$?

From Equation 16.4.2, we have that

$$\int_{-\infty}^{\infty} p_X(x)\,dx = k\int_0^1 \left(x - x^2\right)\,dx = k\left.\frac{x^2}{2} - \frac{x^3}{3}\right|_0^1 = \frac{k}{6}. \tag{16.4.10}$$

Therefore, k must equal 6.

Next, we note that

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} p_X(\xi)\,d\xi. \tag{16.4.11}$$

If $x < 0$, $F_X(x) = 0$. For $0 < x < 1$, then

$$F_X(x) = 6\int_0^x \left(\xi - \xi^2\right)\,d\xi = 6\left.\left(\frac{\xi^2}{2} - \frac{\xi^3}{3}\right)\right|_0^x = 3x^2 - 2x^3. \tag{16.4.12}$$

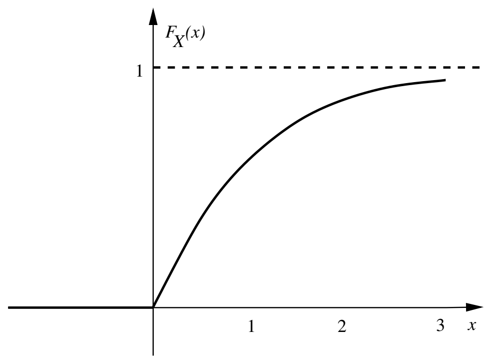


Figure 16.4.2: The cumulative distribution function for an exponential random variable.

Finally, if $x > 1$,

$$F_X(x) = 6 \int_0^1 \left(\xi - \xi^2\right)\, d\xi = 1. \qquad \textbf{(16.4.13)}$$

In summary,

$$F_X(x) = \begin{cases} 0, & 0 \leq x, \\ 3x^2 - 2x^3, & 0 < x \leq 1, \\ 1, & 1 < x. \end{cases} \qquad \textbf{(16.4.14)}$$

Because $P(X \leq x) = F_X(x)$, we have that $P(X < \frac{1}{2}) = \frac{1}{2}$ and $P(X > \frac{1}{2}) = 1 - P(X < \frac{1}{2}) = \frac{1}{2}$. □

• Example 16.4.2: Generating continuous random variables via MATLAB[6]

In the previous section we showed how the MATLAB function `rand` can be used to generate outcomes for a discrete random variable. Similar considerations hold for a continuous random variable.

Consider the exponential random variable X. Its probability density function is

$$p_X(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & 0 < x, \end{cases} \qquad \textbf{(16.4.15)}$$

where $\lambda > 0$. For homework you will show that the corresponding cumulative distribution function is

$$F_X(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-\lambda x}, & 0 < x. \end{cases} \qquad \textbf{(16.4.16)}$$

Figure 16.4.2 illustrates this cumulative density function when $\lambda = 1$. How can we use these results to generate a MATLAB code that produces an exponential random variable?

Recall that both MATLAB function `rand` and the cumulative distribution function produce values that vary between 0 and 1. Given a value from `rand`, we can compute the corresponding $X = x$, which would give the same value from the cumulative distribution function. In short, we are creating random values for the cumulative distribution function and using those values to give the exponential random variable via

$$X = x = -\ln\left(1 - \texttt{rand}\right)/\lambda, \qquad \textbf{(16.4.17)}$$

[6] This technique is known as the inverse transform sampling method. See pages 27–39 in Devroye, L., 1986: *Non-Uniform Random Variable Generation*. Springer-Verlag, 843 pp.

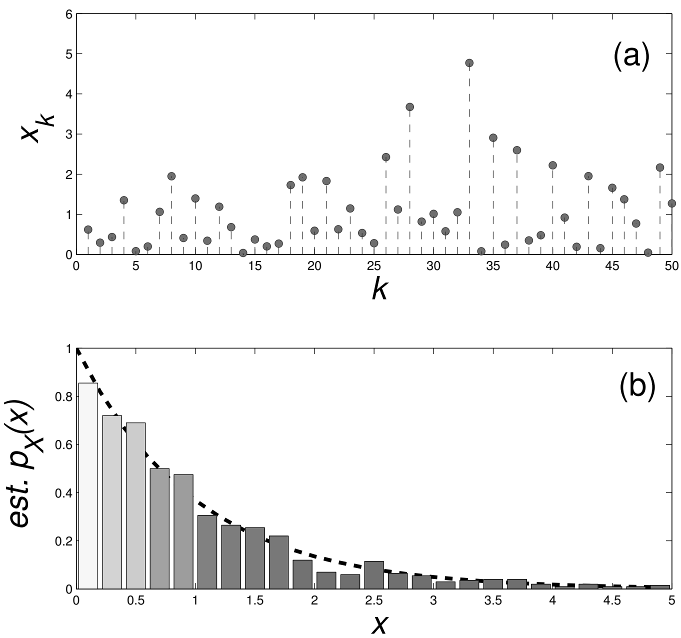


Figure 16.4.3: (a) Outcomes of a numerical experiment to generate an exponential random variable using the MATLAB function `rand`. (b) The function $\hat{p}_X(x)$ given by Equation 16.4.18 as a function of x for an exponential random variable with $M = 1000$. The dashed black line is the exact probability density function.

where we have set $F_X(x) =$ `rand`. Therefore, the MATLAB code to generate exponential random variables for a particular `lambda` is

```
clear;
for i = 1:M
  t = rand(1,1);
  X(i,1) = -log(1-t) / lambda;
end
```

where `M` is the number of experiments that we run. In Figure 16.4.3(a) we illustrates the first 200 outcomes from our numerical experiment to generate an exponential random variable.

To compute the probability density function we use the finite difference approximation of Equation 16.4.1, or

$$\hat{p}(x_0) = \frac{\text{Number of outcomes in } [x_0 - \Delta x/2, x_0 + \Delta x/2]}{M\Delta x}, \qquad \mathbf{(16.4.18)}$$

where Δx is the size of the bins into which we collect the various outcomes. Figure 16.4.3(b) illustrates this numerical estimation of the probability density function in the case of an exponential random variable. The function $\hat{p}_X(x)$ was created from the MATLAB code:

```
clear;
delta_x = 0.2; lambda = 1; M = 1000; % Initialize Δx, λ and M
% sample M outcomes from the uniformly distributed distribution
```

```
t = rand(M,1);
% generate the exponential random variable
x = - log(1-t)/lambda;
% create the various bins [x_0 - Δx/2, x_0 + Δx/2]
bincenters=[delta_x/2:delta_x:5];
bins=length(bincenters); % count the number of bins
% now bin the M outcomes into the various bins
[n,x_out] = hist(x,bincenters);
n = n / (delta_x*M); % compute the probability per bin
bar_h = bar(x_out,n); % create the bar graph
bar_child = get(bar_h,'Children');
set(bar_child,'CData',n);
colormap(Autumn);
```

Problems

1. The probability density function for the exponential random variable is

$$p_X(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & 0 < x, \end{cases}$$

with $\lambda > 0$. Find its cumulative distribution function.

2. Given the probability density function

$$p_X(x) = \begin{cases} kx, & 0 < x < 2, \\ 0, & \text{otherwise}, \end{cases}$$

where k is a constant, (a) compute the value of k, (b) find the cumulative density function $F_X(x)$, and (c) find the $P(1 < X \leq 2)$.

3. Given the probability density function

$$p_X(x) = \begin{cases} k\left(1 - |x|\right), & |x| < 1, \\ 0, & |x| > 1, \end{cases}$$

where k is a constant, (a) compute the value of k and (b) find the cumulative density function $F_X(x)$.

Project: Central Limit Theorem

Consider the sum $S = (X_1 + X_2 + X_3 + \cdots + X_{100})/100$, where X_i is the ith sample from a uniform distribution.

Step 1: Write a MATLAB program to compute the probability density function of S. See Figure 16.4.4.

Step 2: The *central limit theorem* states the distribution of the sum (or average) of a large number of independent, identically distributed random variables will be approximately normal, regardless of the underlying distribution. Do your numerical results agree with this theorem?

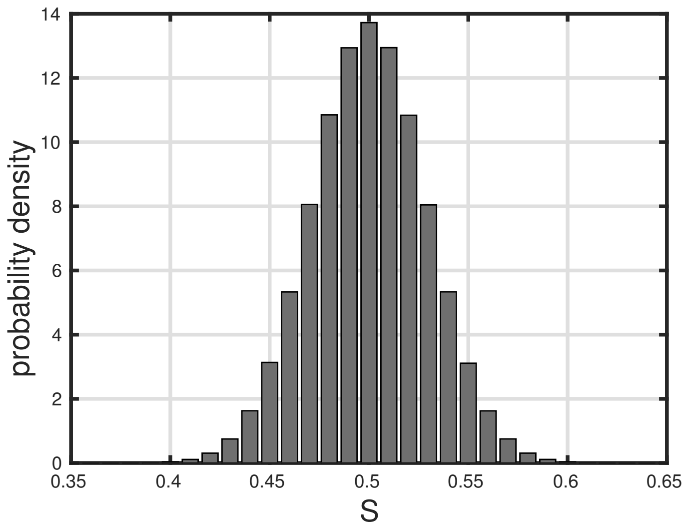


Figure 16.4.4: Computed probability density function for the sum $S = (X_1 + X_2 + X_3 + \cdots + X_{100})/100$, where X_i is the ith sample from a uniform distribution.

16.5 MEAN AND VARIANCE

In the previous two sections we explored the concepts of random variable and distribution. Here we introduce two parameters, *mean* and *variance*, that are useful in characterizing a distribution.

The mean μ_X is defined by

$$\mu_X = E(X) = \begin{cases} \displaystyle\sum_k x_k\, p_X[x_k], & X \text{ discrete}, \\ \displaystyle\int_{-\infty}^{\infty} x\, p_X(x)\, dx, & X \text{ continuous}. \end{cases} \tag{16.5.1}$$

The mean provides the position of the center of the distribution. The operator $E(X)$, which is called the *expectation* of X, gives the average value of X that one should *expect* after many trials.

Two important properties involve the expectation of the sum and product of two random variables X and Y. The first one is

$$E(X + Y) = E(X) + E(Y). \tag{16.5.2}$$

Second, if X and Y are *independent* random variables, then

$$E(XY) = E(X)E(Y). \tag{16.5.3}$$

The proofs can be found elsewhere.[7]

The variance provides the spread of a distribution. It is computed via

$$\sigma_X^2 = \text{Var}(X) = E\{[X - E(X)]^2\}, \tag{16.5.4}$$

[7] For example, Kay, S. M., 2006: *Intuitive Probability and Random Processes Using MATLAB*. Springer, 833 pp. See Sections 7.7 and 12.7.

or

$$\sigma_X^2 = \begin{cases} \displaystyle\sum_k (x_k - \mu_X)^2 p_X[x_k], & X \text{ discrete}, \\ \displaystyle\int_{-\infty}^{\infty} (x - \mu_X)^2 p_X(x)\, dx, & X \text{ continuous}. \end{cases} \quad \textbf{(16.5.5)}$$

If we expand the right side of Equation 16.5.4, an alternative method for finding the variance is

$$\sigma_X^2 = \text{Var}(X) = E(X^2) - [E(X)]^2, \quad \textbf{(16.5.6)}$$

where

$$E(X^n) = \begin{cases} \displaystyle\sum_k x_k^n\, p_X[x_k], & X \text{ discrete}, \\ \displaystyle\int_{-\infty}^{\infty} x^n p_X(x)\, dx, & X \text{ continuous}. \end{cases} \quad \textbf{(16.5.7)}$$

• **Example 16.5.1: Mean and variance of M equally likely outcomes**

Consider the random variable $X = k$ where $k = 1, 2, \ldots, M$. If each event has an equally likely outcome, $p_X[x_k] = 1/M$. Then the expected or average or mean value is

$$\mu_X = \frac{1}{M}\sum_{k=1}^{M} x_k = \frac{M(M+1)}{2M} = \frac{M+1}{2}. \quad \textbf{(16.5.8)}$$

Note that the mean does *not* equal any of the possible values of X. Therefore, the expected value need not equal a value that will be actually observed.

Turning to the variance,

$$\text{Var}(X) = (M+1)\,[(2M+1)/6 - (M+1)/4] \quad \textbf{(16.5.9)}$$
$$= (M+1)\,[4M + 2 - 3M - 3]\,/12 \quad \textbf{(16.5.10)}$$
$$= (M+1)(M-1)/12 = (M^2 - 1)/12, \quad \textbf{(16.5.11)}$$

because

$$E(X^2) = \frac{1}{M}\sum_{k=1}^{M} x_k^2 = \frac{M(M+1)(2M+1)}{6M} = \frac{(M+1)(2M+1)}{6}. \quad \textbf{(16.5.12)}$$

We used Equation 16.5.6 to compute the variance. □

• **Example 16.5.2**

Let us find the mean and variance of the random variable X whose probability density function is

$$p_X(x) = \begin{cases} kx, & 0 < x < 1, \\ 0, & \text{otherwise}. \end{cases} \quad \textbf{(16.5.13)}$$

From Equation 16.5.1, we have that

$$\mu_X = E(X) = \int_0^1 x(kx)\, dx = \left.\frac{kx^3}{3}\right|_0^1 = \frac{k}{3}. \quad \textbf{(16.5.14)}$$

From Equation 16.5.6, the variance of X is

$$\sigma_X^2 = \text{Var}(X) = E(X^2) - [E(X)]^2 = \int_0^1 x^2(kx)\,dx - \frac{k^2}{9} = \left.\frac{kx^4}{4}\right|_0^1 - \frac{k^2}{9} = \frac{k}{4} - \frac{k^2}{9}. \quad \textbf{(16.5.15)}$$

□

• **Example 16.5.3: Characteristic functions**

The *characteristic function* of a random variable is defined by

$$\phi_X(\omega) = E[\exp(i\omega X)]. \quad \textbf{(16.5.16)}$$

If X is a discrete random variable, then

$$\phi_X(\omega) = \sum_{k=-\infty}^{\infty} p_X[x_k]e^{ik\omega}. \quad \textbf{(16.5.17)}$$

On the other hand, if X is a continuous random variable,

$$\phi_X(\omega) = \int_{-\infty}^{\infty} p_X(x)e^{i\omega x}\,dx, \quad \textbf{(16.5.18)}$$

the inverse Fourier transform (times 2π) of the Fourier transform, $p_X(x)$.

Characteristic functions are useful for computing various moments of a random variable via

$$E(X^n) = \frac{1}{i^n} \left.\frac{d^n\varphi_X(\omega)}{d\omega^n}\right|_{\omega=0}. \quad \textbf{(16.5.19)}$$

This follows by taking repeated differentiation of Equation 16.5.16 and then evaluating the differentiation at $\omega = 0$.

Consider, for example, the exponential probability density function $p_X(x) = \lambda e^{-\lambda x}$ with $x, \lambda > 0$. A straightforward calculation gives

$$\phi_X(\omega) = \frac{\lambda}{\lambda - \omega i}. \quad \textbf{(16.5.20)}$$

Substituting Equation 16.5.20 into Equation 16.5.19 yields

$$E(X^n) = \frac{n!}{\lambda^n}. \quad \textbf{(16.5.21)}$$

In particular,

$$E(X) = \frac{1}{\lambda} \qquad \text{and} \qquad E(X^2) = \frac{2}{\lambda^2}. \quad \textbf{(16.5.22)}$$

Consequently, $\mu_X = 1/\lambda$ and $\text{Var}(X) = E(X^2) - \mu_X^2 = 1/\lambda^2$. □

• **Example 16.5.4: Characteristic function for a Gaussian distribution**

Let us find the characteristic function for the Gaussian distribution and then use that characteristic function to compute the mean and variance.

Because

$$p_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \tag{16.5.23}$$

the characteristic function equals

$$\phi_X(\omega) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)+i\omega x}\, dx \tag{16.5.24}$$

$$= e^{i\omega\mu-\sigma^2\omega^2/2} \left\{ \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\mu-i\omega\sigma^2)^2}{2\sigma^2} \right] dx \right\} \tag{16.5.25}$$

$$= e^{i\omega\mu-\sigma^2\omega^2/2} \tag{16.5.26}$$

because the quantity within the wavy brackets equals one.

Given this characteristic function, Equation 16.5.26, we have that

$$\phi_X'(\omega) = (i\mu - \sigma^2\omega)e^{i\omega\mu-\sigma^2\omega^2/2}. \tag{16.5.27}$$

Therefore, $\phi_X'(0) = i\mu$ and from Equation 16.5.19, $\mu_X = E(X) = \mu$. Furthermore,

$$\phi_X''(\omega) = (i\mu - \sigma^2\omega)^2 e^{i\omega\mu-\sigma^2\omega^2/2} - \sigma^2 e^{i\omega\mu-\sigma^2\omega^2/2}. \tag{16.5.28}$$

Consequently, $\phi_X''(0) = -\mu^2 - \sigma^2$ and $\mathrm{Var}(X) = E(X^2) - \mu_X^2 = \sigma^2$.

Problems

1. Let $X(s)$ denote a discrete random variable associated with a fair coin toss. Then

$$X(s) = \begin{cases} 0, & s = \text{tail}, \\ 1, & s = \text{head}. \end{cases}$$

Find the expected value and variance of this random variable.

2. The geometric random variable X has the probability mass function:

$$p_X[x_k] = P(X = k) = p(1-p)^{k-1}, \qquad k = 1, 2, 3, \ldots.$$

Find its mean and variance. Hint:

$$\sum_{k=1}^{\infty} kr^{k-1} = \frac{1}{(1-r)^2}, \qquad \sum_{k=2}^{\infty} k(k-1)r^{k-2} = \frac{2}{(1-r)^3}, \quad |r| < 1,$$

and $E(X^2) = E[X(X-1)] + E(X)$.

3. Given

$$p_X(x) = \begin{cases} kx(2-x) & 0 < x < 2, \\ 0, & \text{otherwise}, \end{cases}$$

(a) find k and (b) its mean and variance.

4. Given the probability density

$$p_X(x) = (a^2 - x^2)^{\nu-\frac{1}{2}}, \qquad \nu > -\tfrac{1}{2},$$

find its characteristic function using integral tables.

For the following distributions, first find their characteristic functions. Then compute the mean and variance using Equation 16.5.19.

5. Binomial distribution:

$$p_X[x_k] = \binom{n}{k} p^k q^{n-k}, \qquad 0 < p < 1,$$

where $q = 1 - p$. Hint: Use the binomial theorem to simplify Equation 16.5.17.

6. Poisson distribution:

$$p_X[x_k] = e^{-\lambda} \frac{\lambda^k}{k!}, \qquad 0 < \lambda.$$

7. Geometric distribution:

$$p_X[x_k] = q^k p, \qquad 0 < p < 1,$$

where $q = 1 - p$.

8. Uniform distribution:

$$p_X(x) = \frac{H(x-a) - H(x-b)}{b-a}, \qquad b > a > 0.$$

Project: MATLAB's Intrinsic Function `mean` and `var`

MATLAB has the special commands `mean` and `var` to compute the mean and variance, respectively, of the random variable X. Use the MATLAB command `randn` to create a random variable `X(n)` of length `N`. Then, find the mean and variance of `X(n)`. How do these parameters vary with `N`?

Project: Monte Carlo Integration and Importance Sampling

Consider the integral $I = \int_0^1 \sqrt{1-x^2}\, dx = \pi/4$. If we were to compute it numerically by the conventional midpoint rule, the approximate value is given by

$$I_N = \frac{1}{N} \sum_{n=1}^{N} f(x_n), \qquad (1)$$

where $f(x) = \sqrt{1-x^2}$ and $x_n = (n - 1/2)/N$. For $N = 10$, 50, 100, and 500, the absolute value of the relative error is 2.7×10^{-3}, 2.4×10^{-4}, 8.6×10^{-5}, and 7.7×10^{-6}, respectively.

Monte Carlo integration is a simple alternative method for doing the numerical integration using random sampling. It is a particularly powerful technique for approximating complicated integrals. Here you will explore a simple one-dimensional version of this scheme.

Consider the random variable:

$$I_M = \frac{1}{M} \sum_{m=1}^{M} f(x_m), \qquad (2)$$

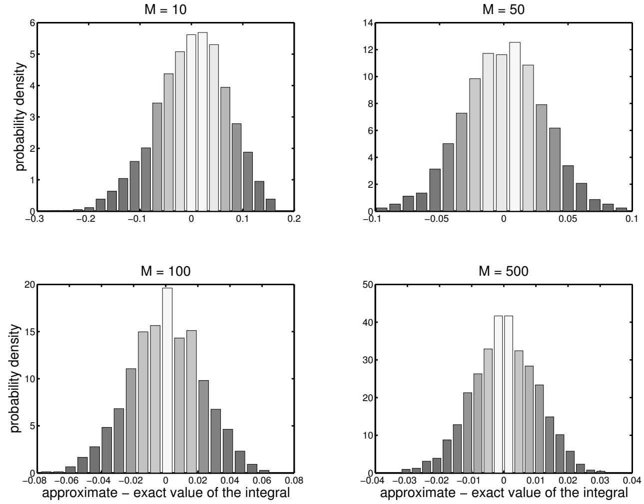


Figure 16.5.1: The probability density function arising from using Monto Carlo integration to compute $\int_0^1 \sqrt{1-x^2}\, dx$ for various values of M.

where x_m is the mth sample point taken from the uniform distribution. I_M is a random variable because it is a function of the random variable x_m. Therefore,

$$E(I_M) = \frac{1}{M}\sum_{m=1}^{M} E[f(x_m)] = \frac{1}{M}\sum_{m=1}^{M}\int_0^1 f(x)p(x)\, dx$$
$$= \frac{1}{M}\sum_{m=1}^{M}\int_0^1 f(x)\, dx = \int_0^1 f(x)\, dx = I,$$

because $p(x)$, the probability of the uniform distribution, equals 1. Furthermore, as we increase the number of samples M, I_M approaches I. By the *strong law of large numbers*, this limit is guaranteed to converge to the exact solution: $P(\lim_{M\to\infty} I_M - I) = 1$. Equation (2) is *not* the midpoint rule because the uniform grid x_n has been replaced by randomly spaced grid points.

Step 1: Write a MATLAB program that computes I_M for various values of M when x_m is selected from a uniform distribution. By running your code thousands of times, find the probability density as a function of the difference between I_M and I. Compute the mean and variance of I_M. How does the variance vary with M? See Figure 16.5.1.

The reason why standard Monte Carlo integration is not particularly good is the fact that we used a uniform distribution. A better idea would be to sample from regions where the integrand is larger. This is the essence of the concept of *importance sampling*: That certain values of the input random variable x_m in a simulation have more impact on the parameters being estimated than others.

We begin by noting that

$$I = \int_0^1 f(x)\, dx = \int_0^1 \frac{f(x)}{p_1(x)} p_1(x)\, dx,$$

where $p_1(x)$ is a new probability density function that replaces the uniform probability distribution and is relatively larger when $f(x)$ is larger and relatively smaller when $f(x)$ is smaller.

The question now becomes how to compute $p_1(x)$. We shall use the VEGAS algorithm, which constructs $p_1(x)$ by sampling $f(x)$ K times, where $K < M$. Within each kth subinterval we assume that there are M/K uniformly distributed points. Therefore,

$$p_1(x_m) = \frac{K\, f(s_m)}{\sum_{k=1}^{K} f(s_k)},$$

where s_k is the center point of the kth subinterval within which the mth point is located. For each m, we must find x_m. This is done in two steps: First we randomly choose the kth subinterval using a uniform distribution. Then we randomly choose the point x_m within that subinterval using a uniform distribution. Therefore, our modified integration scheme becomes

$$I_M = \frac{1}{M} \sum_{m=1}^{M} \frac{f(x_m)}{p_1(x_m)}. \qquad (\mathbf{3})$$

Now,

$$E(I_M) = \frac{1}{M} \sum_{m=1}^{M} E\left[\frac{f(x_m)}{p_1(x_m)}\right] = \frac{1}{M} \sum_{m=1}^{M} \int_0^1 \frac{f(x)}{p_1(x)} p_1(x) p_2(x)\, dx$$
$$= \frac{1}{M} \sum_{m=1}^{M} \int_0^1 f(x)\, dx = \int_0^1 f(x)\, dx = I,$$

because $p_2(x) = 1$.

Step 2: Write a MATLAB program that computes I_M for various values of K for a fixed value of M. Recall that you must first select the subdivision using the *MATLAB* function `rand` and then the value of x_m within the subdivision using a uniform distribution. By running your code thousands of times, find the probability density as a function of the difference between I_M and I. Compute the mean and variance of I_M. How does the variance vary with M? See Figure 16.5.2.

16.6 SOME COMMONLY USED DISTRIBUTIONS

In the previous sections we introduced the concept of probability distributions and their description via mean and variance. In this section we focus on some special distributions, both discrete and continuous, that appear often in engineering.

Bernoulli distribution

Consider an experiment where the outcome can be classified as either a success or failure. The probability of a success is p and the probability of a failure is $1-p$. Then these "Bernoulli trials" have a random variable X associated with them where the probability mass function is given by

$$p_X[x_k] = P(X = k) = p^k (1-p)^{1-k}, \qquad k = 0, 1, \qquad (\mathbf{16.6.1})$$

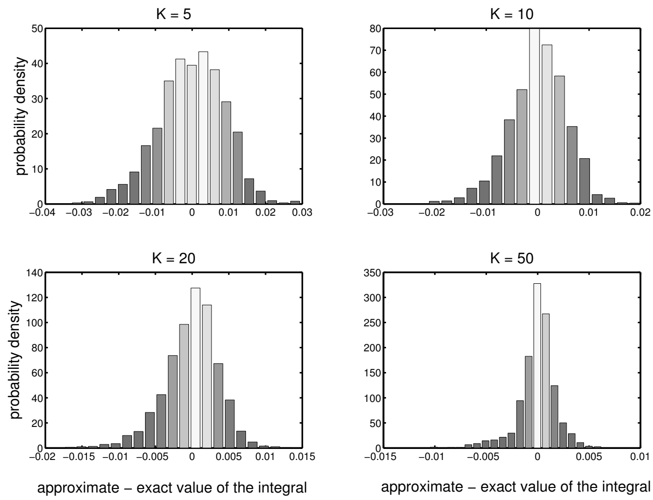


Figure 16.5.2: The probability density function arising from using importance sampling with Monto Carlo integration to compute $\int_0^1 \sqrt{1-x^2}\, dx$ for various values of K and $M = 100$.

where $0 \le p \le 1$. From Equation 16.3.12 the cumulative density function of the Bernoulli random variable X is

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1-p, & 0 \le x < 1, \\ 1, & 1 \le x. \end{cases} \tag{16.6.2}$$

The mean and variance of the Bernoulli random variable X are

$$\mu_X = E(X) = p, \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = p(1-p). \tag{16.6.3}$$

• Example 16.6.1

A simple pass and fail process is taking a final exam, which can be modeled by a Bernoulli distribution. Suppose a class passed a final exam with the probability of 0.75. If X denotes the random variable that someone passed the exam, then

$$E(X) = p = 0.75, \qquad \text{and} \qquad \text{Var}(X) = p(1-p) = (0.75)(0.25) = 0.1875. \tag{16.6.4}$$

□

Geometric distribution

Consider again an experiment where we either have success with probability p or failure with probability $1-p$. This experiment is repeated until the first success occurs. Let random variable X denote the trial number on which this first success occurs. Its probability mass function is

$$p_X[x_k] = P(X = k) = p(1-p)^{k-1}, \qquad k = 1, 2, 3, \ldots. \tag{16.6.5}$$

From Equation 16.3.12 the cumulative density function of this geometric random variable X is

$$F_X(x) = P(X \le x) = 1 - (1-p)^k. \tag{16.6.6}$$

The mean and variance of the geometric random variable X are

$$\mu_X = E(X) = \frac{1}{p}, \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = \frac{1-p}{p^2}. \tag{16.6.7}$$

• Example 16.6.2

A particle within an accelerator has the probability 0.01 of hitting a target material. (a) What is the probability that the first particle to hit the target is the 50th? (b) What is the probability that the target will be hit by *any* particle?

$$P(\text{first particle to hit is the 50th}) = 0.01(0.99)^{49} = 0.0061. \tag{16.6.8}$$

$$P(\text{target hit by any of first 50th particles}) = \sum_{n=1}^{50} 0.01(0.99)^{n-1} = 0.3950. \tag{16.6.9}$$

□

• Example 16.6.3

The police ticket 5% of parked cars. Assuming that the cars are ticketed independently, find the probability of 1 ticket on a block with 7 parked cars.

Each car is a Bernoulli trial with $P(\text{ticket}) = 0.05$. Therefore,

$$P(1 \text{ ticket on block}) = P(1 \text{ ticket in 7 trials}) = \binom{7}{1}(0.95)^6(0.05) = 0.2573. \tag{16.6.10}$$

□

Binomial distribution

Consider now an experiment in which n independent Bernoulli trials are performed and X represents the number of successes that occur in the n trials. In this case the random variable X is called *binomial* with parameters (n, p) with a probability mass function given by

$$p_X[x_k] = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n, \tag{16.6.11}$$

where $0 \le p \le 1$, and

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \tag{16.6.12}$$

the *binomial coefficient*. The term p^k arises from the k successes while $(1-p)^{n-k}$ is due to the failures. The binomial coefficient gives the number of ways that we pick those k successes from the n trials.

The corresponding cumulative density function of X is

$$F_X(x) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}, \qquad n \le x < n+1. \tag{16.6.13}$$

The mean and variance of the binomial random variable X are

$$\mu_X = E(X) = np, \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = np(1-p). \tag{16.6.14}$$

A Bernoulli random variable is the same as a binomial random variable when the parameters are $(1, n)$.

• Example 16.6.4

Let us find the probability of rolling the same side of a die (say, the side with N dots on it) at least 3 times when a fair die is rolled 4 times.

During our 4 tosses, we could obtain no rolls with N dots on the side ($k=0$), one roll with N dots ($k=1$), two rolls with N dots ($k=2$), three rolls with N dots ($k=3$), or four rolls with N dots ($k=4$). If we define A as the event of rolling a die so that the side with N dots appears *at least* three times, then we must add the probabilities for $k=3$ and $k=4$. Therefore,

$$P(A) = p_X[x_3] + p_X[x_4] = \binom{4}{3} p^3(1-p)^1 + \binom{4}{4} p^4(1-p)^0 \tag{16.6.15}$$

$$= \frac{4!}{3!1!}p^3(1-p)^1 + \frac{4!}{4!0!}p^4(1-p)^0 = 0.0162 \tag{16.6.16}$$

because $p = \frac{1}{6}$. □

• Example 16.6.5

If 10 random binary digits are transmitted, what is the probability that *more* than seven 1's are included among them?

Let X denote the number of 1's among the 10 digits. Then

$$P(X > 7) = P(X=8) + P(X=9) + P(X=10) = p_X[x_8] + p_X[x_9] + p_X[x_{10}] \tag{16.6.17}$$

$$= \binom{10}{8}\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^2 + \binom{10}{9}\left(\frac{1}{2}\right)^9\left(\frac{1}{2}\right)^1 + \binom{10}{10}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^0 \tag{16.6.18}$$

$$= (45 + 10 + 1)\left(\frac{1}{2}\right)^{10} = \frac{56}{1024}. \tag{16.6.19}$$

□

Poisson distribution

The Poisson probability distribution arises as an approximation for the binomial distribution as $n \to \infty$ and $p \to 0$ such that np remains finite. To see this, let us rewrite the binomial distribution as follows:

$$P(X=k) = \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}\frac{\lambda^n}{n!}\frac{(1-\lambda/n)^n}{(1-\lambda/n)^k}, \tag{16.6.20}$$

if $\lambda = np$. For finite λ,

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^k \to 1, \qquad \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}, \tag{16.6.21}$$

and

$$\lim_{n\to\infty}\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \to 1. \tag{16.6.22}$$

Therefore, for large n, small p and moderate λ, we can approximate the binomial distribution by the Poisson distribution:

$$p_X[x_k] = P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k=0,1,\ldots. \tag{16.6.23}$$

The corresponding cumulative density function of X is

$$F_X(x) = e^{-\lambda}\sum_{k=0}^{n}\frac{\lambda^k}{k!}, \qquad n \le x < n+1. \tag{16.6.24}$$

The mean and variance of the Poisson random variable X are

$$\mu_X = E(X) = \lambda, \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = \lambda. \tag{16.6.25}$$

In addition to this approximation, the Poisson distribution is the probability distribution for a Poisson process. But that has to wait for the next chapter.

• Example 16.6.6

Consider a student union on a campus. On average 3 persons enter the union per minute. What is the probability that, during any given minute, 3 or more persons will enter the union?

To make use of Poisson's distribution to solve this problem, we must have both a large n and a small p with the average $\lambda = np = 3$. Therefore, we divide time into a large number of small intervals so that n is large while the probability that someone will enter the union is small. Assuming independence of events, we have a binomial distribution with large n. Let A denote the event that 3 or more persons will enter the union, then

$$P(\overline{A}) = p_X[0] + p_X[1] + p_X[2] = e^{-3}\left(\frac{3^0}{0!}+\frac{3^1}{1!}+\frac{3^2}{2!}\right) = 0.423. \tag{16.6.26}$$

Therefore, $P(A) = 1 - P(\overline{A}) = 0.577$. □

Uniform distribution

The continuous random variable X is called *uniform* if its probability density function is

$$p_X(x) = \begin{cases} 1/(b-a), & a < x < b, \\ 0, & \text{otherwise.} \end{cases} \tag{16.6.27}$$

The corresponding cumulative density function of X is

$$F_X(x) = \begin{cases} 0, & x \le a, \\ (x-a)/(b-a), & a < x < b, \\ 1, & b \le x. \end{cases} \tag{16.6.28}$$

The mean and variance of a uniform random variable X are

$$\mu_X = E(X) = \tfrac{1}{2}(a+b), \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = \frac{(b-a)^2}{12}. \tag{16.6.29}$$

Uniform distributions are used when we have no prior knowledge of the actual probability density function and all continuous values in some range appear equally likely.

Exponential distribution

The continuous random variable X is called *exponential* with parameter $\lambda > 0$ if its probability density function is

$$p_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x < 0. \end{cases} \tag{16.6.30}$$

The corresponding cumulative density function of X is

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0, \\ 0, & x < 0. \end{cases} \tag{16.6.31}$$

The mean and variance of an exponential random variable X are

$$\mu_X = E(X) = 1/\lambda, \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = 1/\lambda^2. \tag{16.6.32}$$

This distribution has the interesting property that is "memoryless." By memoryless, we mean that for a nonnegative random variable X, then

$$P(X > s + t | X > t) = P(X > s), \tag{16.6.33}$$

where $x, t \ge 0$. For example, if the lifetime of a light bulb is exponentially distributed, then the light bulb that has been in use for some hours is as good as a new light bulb with regard to the amount of time remaining until it fails.

To prove this, from Equation 16.2.4, Equation 16.6.33 becomes

$$\frac{P(X > s + t \text{ and } X > t)}{P(X > t)} = P(X > s), \tag{16.6.34}$$

or

$$P(X > s + t \text{ and } X > t) = P(X > t)P(X > s), \tag{16.6.35}$$

since $P(X > s + t \text{ and } X > t) = P(X > s + t)$. Now, because

$$P(X > s + t) = 1 - \left[1 - e^{-\lambda(s+t)}\right] = e^{-\lambda(s+t)}, \tag{16.6.36}$$

$$P(X > s) = 1 - \left(1 - e^{-\lambda s}\right) = e^{-\lambda s}, \tag{16.6.37}$$

and

$$P(X > t) = 1 - \left(1 - e^{-\lambda t}\right) = e^{-\lambda t}. \tag{16.6.38}$$

Therefore, Equation 16.6.35 is satisfied and X is memoryless.

• Example 16.6.7

A component in an electrical circuit has an exponentially distributed failure time with a mean of 1000 hours. Calculate the time so that the probability of the time to failure is less than 10^{-3}.

Let the exponential random variable $X = k$ have the units of hours. Then $\lambda = 10^{-3}$. From the definition of the cumulative density function,

$$F_X(x_t) = P(X \le x_t) = 0.001, \qquad \text{and} \qquad 1 - \exp(-\lambda x_t) = 0.001. \tag{16.6.39}$$

Solving for x_t,

$$x_t = -\ln(0.999)/\lambda = 1. \tag{16.6.40}$$

□

• Example 16.6.8

A computer contains a certain component whose time (in years) to failure is given by the random variable T distributed exponentially with $\lambda = 1/5$. If 5 of these components are installed in different computers, what is the probability that at least 2 of them will still work at the end of 8 years?

The probability that a component will last 8 years or longer is

$$P(T > 8) = e^{-8/5} = 0.2019, \tag{16.6.41}$$

because $\lambda = 1/5$.

Let X denote the number of components functioning after 8 years. Then,

$$P(X \ge 2) = 1 - P(X = 0) - P(X = 1) \tag{16.6.42}$$

$$= 1 - \binom{5}{0}(0.2019)^0(0.7981)^5 - \binom{5}{1}(0.2019)^1(0.7981)^4 \tag{16.6.43}$$

$$= 0.2666. \tag{16.6.44}$$

□

Normal (or Gaussian) distribution

The normal distribution is the most important continuous distribution. It occurs in many applications and plays a key role in the study of random phenomena in nature.

A random variable X is called a *normal* random variable if its probability density function is

$$p_X(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi}\,\sigma}, \tag{16.6.45}$$

where the mean and variance of a normal random variable X are

$$\mu_X = E(X) = \mu, \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = \sigma^2. \tag{16.6.46}$$

The distribution is symmetric with respect to $x = \mu$ and its shape is sometimes called "bell shaped." For small σ^2 we obtain a high peak and steep slope while with increasing σ^2 the curve becomes flatter and flatter.

The corresponding cumulative density function of X is

$$F_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-(\xi-\mu)^2/(2\sigma^2)}\, d\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-\xi^2/2}\, d\xi. \tag{16.6.47}$$

The integral in Equation 16.6.46 must be evaluated numerically. It is convenient to introduce the *probability integral*:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\xi^2/2}\, d\xi. \tag{16.6.48}$$

Note that $\Phi(-z) = 1 - \Phi(z)$. Therefore,

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right). \tag{16.6.49}$$

and

$$P(a < X \le b) = F_X(b) - F_X(a). \tag{16.6.50}$$

Consider now the intervals consisting of one σ, two σ, and three σ around the mean μ. Then, from Equation 16.6.50,

$$P(\mu - \sigma < X \le \mu + \sigma) = 0.68, \tag{16.6.51}$$

$$P(\mu - 2\sigma < X \le \mu + 2\sigma) = 0.955, \tag{16.6.52}$$

and

$$P(\mu - 3\sigma < X \le \mu + 3\sigma) = 0.997. \tag{16.6.53}$$

Therefore, approximately $\frac{2}{3}$ of the values will be distributed between $\mu - \sigma$ and $\mu + \sigma$, approximately 95% of the values will be distributed between $\mu - 2\sigma$ and $\mu + 2\sigma$, and almost all values will be distributed between $\mu - 3\sigma$ and $\mu + 3\sigma$. For most uses, then, all values will lie between $\mu - 3\sigma$ and $\mu + 3\sigma$, the so-called "three-sigma limits."

As stated earlier, the mean and variance of a normal random variable X are

$$\mu_X = E(X) = \mu, \qquad \text{and} \qquad \sigma_X^2 = \text{Var}(X) = \sigma^2. \tag{16.6.54}$$

The notation $N(\mu; \sigma)$ commonly denotes that X is normal with mean μ and variance σ^2. The special case of a normal random variable Z with zero mean and unit variance, $N(0, 1)$, is called a *standard normal random variable*.

Problems

1. Four coins are tossed simultaneously. Find the probability function for the random variable X that gives the number of heads. Then compute the probabilities of (a) obtaining no heads, (b) exactly one head, (c) at least one head, and (d) not less than four heads.

2. A binary source generates the digits 1 and 0 randomly with equal probability. (a) What is the probability that three 1's and three 0's will occur in a six-digit sequence? (b) What is the probability that *at least* three 1's will occur in a six-digit sequence?

3. Show that the probability of exactly n heads in $2n$ tosses of a fair coin is

$$p_X[x_n] = \frac{1 \cdot 3 \cdot 5 \cdots 2n-1}{2 \cdot 4 \cdot 6 \cdots 2n}.$$

4. If your cell phone rings, on average, 3 times between noon and 3 P.M., what is the probability that during that time period you will receive (a) no calls, (b) 6 or more calls, and (c) not more than 2 calls? Assume that the probability is given by a Poisson distribution.

5. A company sells blank DVDs in packages of 10. If the probability of a defective DVD is 0.001, (a) what is the probability that a package contains a defective DVD? (b) what is the probability that a package has two or more defective DVDs?

6. A school plans to offer a course on probability in a classroom that contains 20 seats. From experience they know that 95% of the students who enroll actually show up. If the school allows 22 students to enroll before the session is closed, what is the probability of the class being oversubscribed?

7. The lifetime (in hours) of a certain electronic device is a random variable T having a probability density function $p_T(t) = 100H(t-100)/t^2$. What is the probability that exactly 3 of 5 such devices must be replaced within the first 150 hours of operation? Assume that the events that the ith device must be replaced within this time are independent.

16.7 JOINT DISTRIBUTIONS

In the previous sections we introduced distributions that depended upon a single random variable. Here we generalize these techniques for two random variables. The range of the two-dimensional random variable (X,Y) is $R_{XY} = \{(x,y); \xi \in S \text{ and } X(\xi) = x, Y(\xi) = y\}$.

Discrete joint distribution

Let X and Y denote two *discrete* random variables defined on the same sample space (jointly distributed). The function $p_{XY}[x_i, y_j] = P[X = x_i, Y = y_j]$ is the *joint probability mass function* of X and Y. As one might expect, $p_{XY}[x_i, y_j] \geq 0$.

Let the sets of possible values of X and Y be A and B. If $x_i \notin A$ or $y_j \notin B$, then $p_{XY}[x_i, y_j] = 0$. Furthermore,

$$\sum_{x_i \in A, y_j \in B} p_{XY}[x_i, y_j] = 1. \qquad \textbf{(16.7.1)}$$

The *marginal probability functions* of X and Y are defined by

$$p_X[x_i] = \sum_{y_j \in B} p_{XY}[x_i, y_j], \qquad \textbf{(16.7.2)}$$

and

$$p_Y[y_j] = \sum_{x_i \in A} p_{XY}[x_i, y_j]. \tag{16.7.3}$$

If X and Y are independent random variables, then $p_{XY}[x_i, y_j] = p_X[x_i] \cdot p_Y[y_j]$.

• Example 16.7.1

A joint probability mass function is given by

$$p_{XY}[x_i, y_j] = \begin{cases} k(x_i + 2y_j), & x_i = 1,2,3, y_j = 1,2; \\ 0, & \text{otherwise.} \end{cases} \tag{16.7.4}$$

Let us find the value of k, $p_X[x_i]$, and $p_Y[y_j]$.

From Equation 16.7.1, we have that

$$k \sum_{x_i=1}^{3} \sum_{y_j=1}^{2} (x_i + 2y_j) = 1, \tag{16.7.5}$$

or

$$k\big[(1+2) + (1+4) + (2+2) + (2+4) + (3+2) + (3+4)\big] = 1. \tag{16.7.6}$$

Therefore, $k = 1/30$.

Turning to $p_X[x_i]$ and $p_Y[y_j]$,

$$p_X[x_i] = k \sum_{y_j=1}^{2} (x_i + 2y_j) = k(x_i + 2) + k(x_i + 4) = k(2x_i + 6) = (x_i + 3)/15, \tag{16.7.7}$$

where $x_i = 1, 2, 3$, and

$$p_Y[y_j] = k \sum_{x_i=1}^{3} (x_i + 2y_j) = k(1 + 2y_j) + k(2 + 2y_j) + k(3 + 2y_j) = k(6 + 6y_j) = (1 + y_j)/5, \tag{16.7.8}$$

where $y_j = 1, 2$. □

• Example 16.7.2

Consider an urn that contains 1 red ball, 2 blue balls, and 2 green balls. Let (X, Y) be a bivariate random variable where X and Y denote the number of red and blue balls, respectively, chosen from the urn. There are 18 possible ways that three balls can be drawn from the urn: *rbb*, *rbg*, *rgb*, *rgg*, *brb*, *brg*, *bbr*, *bbg*, *bgr*, *bgb*, *bgg*, *grb*, *grg*, *gbr*, *gbb*, *gbg*, *ggr*, and *ggb*.

The range of X and Y in the present problem is $R_{XY} = \{(0,1), (0,2), (1,0), (1,1), (1,2)\}$. The joint probability mass function of (X, Y) is given by $p_{XY}[x_i, y_j] = P(X = i, Y = j)$, where $x_i = 0, 1$ and $y_j = 0, 1, 2$. From our list of possible drawings, we find that $p_{XY}[0,0] = 0$, $p_{XY}[0,1] = 1/6$, $p_{XY}[0,2] = 1/6$, $p_{XY}[1,0] = 1/6$, $p_{XY}[1,1] = 1/3$, and $p_{XY}[1,2] = 1/6$. Note that all of these probabilities sum to one.

Given these probabilities, the marginal probabilities are $p_X[0] = 1/3$, $p_X[1] = 2/3$, $p_Y[0] = 1/3$, $p_Y[1] = 1/2$, and $p_Y[2] = 1/3$. Because $p_{XY}[0,0] \neq p_X[0] p_Y[0]$, X and Y are not independent variables. □

• **Example 16.7.3**

Consider a community where 50% of the families have a pet. Of these families, 60% have one pet, 30% have 2 pets, and 10% have 3 pets. Furthermore, each pet is equally likely (and independently) to be a male or female. If a family is chosen at random from the community, then we want to compute the joint probability that his family has M male pets and F female pets.

These probabilities are as follows:

$$P\{F=0, M=0\} = P\{\text{no pets}\} = 0.5, \tag{16.7.9}$$

$$\begin{aligned} P\{F=1, M=0\} &= P\{\text{1 female and total of 1 pet}\} &(16.7.10)\\ &= P\{\text{1 pet}\}P\{\text{1 female}|\text{1 pet}\} &(16.7.11)\\ &= (0.5)(0.6)\times\frac{1}{2} = 0.15, &(16.7.12)\end{aligned}$$

$$\begin{aligned} P\{F=2, M=0\} &= P\{\text{2 females and total of 2 pets}\} &(16.7.13)\\ &= P\{\text{2 pets}\}P\{\text{2 females}|\text{2 pets}\} &(16.7.14)\\ &= (0.5)(0.3)\times\left(\frac{1}{2}\right)^2 = 0.0375, &(16.7.15)\end{aligned}$$

and

$$\begin{aligned} P\{F=3, M=0\} &= P\{\text{3 females and total of 3 pets}\} &(16.7.16)\\ &= P\{\text{3 pets}\}P\{\text{3 females}|\text{3 pets}\} &(16.7.17)\\ &= (0.5)(0.1)\times\left(\frac{1}{2}\right)^3 = 0.00625. &(16.7.18)\end{aligned}$$

The remaining probabilities can be obtained in a similar manner. □

Continuous joint distribution

Let us now turn to the case when we have two continuous random variables. In analog with the definition given in Section 16.4, we define the two-dimensional probability density $p_{XY}(x,y)$ by

$$P(x < X \le x+dx, y < Y \le y+dy) = p_{XY}(x,y)\,dx\,dy. \tag{16.7.19}$$

Here, the comma in the probability parentheses means "and also."

Repeating the same analysis as in Section 16.4, we find that $p_{XY}(x,y)$ must be a single-valued function with $p_{XY}(x,y) \ge 0$, and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{XY}(x,y)\,dx\,dy = 1. \tag{16.7.20}$$

The joint distribution function of X and Y is

$$F_{XY}(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p_{XY}(\xi,\eta)\,d\xi\,d\eta. \tag{16.7.21}$$

Therefore,

$$P(a < X \le b, c < Y \le d) = \int_{a}^{b} \int_{c}^{d} p_{XY}(\xi,\eta)\,d\xi\,d\eta. \tag{16.7.22}$$

The *marginal probability density functions* are defined by

$$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x,y)\,dy, \quad \text{and} \quad p_Y(y) = \int_{-\infty}^{\infty} p_{XY}(x,y)\,dx. \tag{16.7.23}$$

An important distinction exists upon whether the random variables are independent or not. Two variables X and Y are *independent* if and only if

$$p_{XY}(x,y) = p_X(x)p_Y(y), \tag{16.7.24}$$

and conversely.

• Example 16.7.4

The joint probability density function of bivariate random variables (X,Y) is

$$p_{XY}(x,y) = \begin{cases} kxy, & 0 < y < x < 1, \\ 0, & \text{otherwise}, \end{cases} \tag{16.7.25}$$

where k is a constant. (a) Find the value of k. (b) Are X and Y independent?

The range R_{XY} for this problem is a right triangle with its sides given by $x = 1$, $y = 0$, and $y = x$. From Equation 16.7.20,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{XY}(x,y)\,dx\,dy = k\int_0^1 x\left[\int_0^x y\,dy\right]\,dx = k\int_0^1 x\left.\frac{y^2}{2}\right|_0^x\,dx \tag{16.7.26}$$

$$= \frac{k}{2}\int_0^1 x^3\,dx = \frac{k}{8}\,x^4\big|_0^1 = \frac{k}{8}. \tag{16.7.27}$$

Therefore, $k = 8$.

To check for independence we must first compute $p_X(x)$ and $p_Y(y)$. From Equation 16.7.23 and holding x constant,

$$p_X(x) = 8x\int_0^x y\,dy = 4x^3, \qquad 0 < x < 1; \tag{16.7.28}$$

$p_X(x) = 0$ otherwise. From Equation 16.7.23 and holding y constant,

$$p_Y(y) = 8y\int_y^1 x\,dx = 4y(1-y^2), \qquad 0 < y < 1. \tag{16.7.29}$$

Because $p_{XY}(x,y) \ne p_X(x)p_Y(y)$, X and Y are not independent. □

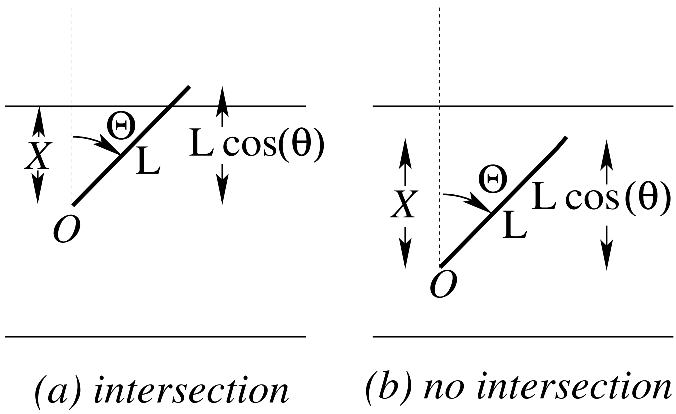


Figure 16.7.1: Schematic of Buffon's needle problem showing the random variables X and Θ.

• **Example 16.7.5: Buffon's needle problem**

A classic application of joint probability distributions is the solution of Buffon's needle problem:[8] Consider an infinite plane with an infinite series of parallel lines spaced a unit distance apart. A needle of length $L < 1$ is thrown upward and we want to compute the probability that the stick will land so that it intersects one of these lines. See Figure 16.7.1.

There are two random variables that determine the needle's orientation: X, the distance from the lower end O of the needle to the nearest line above and Θ, the angle from the vertical to the needle. Of course, we assume that the position where the needle lands is random; otherwise, it would not be a probability problem.

Let us define X first. Its possible values lie between 0 and 1. Second, X is uniformly distributed on $(0, 1)$ with the probability density

$$p_X(x) = \begin{cases} 1, & 0 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases} \tag{16.7.30}$$

Turning to Θ, its value lies between $-\pi/2$ to $\pi/2$ and is uniformly distributed between these values. Therefore, the probability density is

$$p_\Theta(\theta) = \begin{cases} 1/\pi, & -\pi/2 < \theta < \pi/2, \\ 0, & \text{otherwise.} \end{cases} \tag{16.7.31}$$

The probability p that we seek is

$$p = P\{\text{needle intersects line}\} = P\{X < L\cos(\Theta)\}. \tag{16.7.32}$$

Because X and Θ are independent, their joint density equals the product of the densities for X and Θ: $p_{X\Theta}(x, \theta) = p_X(x)p_\Theta(\theta)$.

The final challenge is to use $p_{X\Theta}(x, \theta)$ to compute p. In Section 16.2 we gave a geometric definition of probability. The area of the sample space is π because it consists of a rectangle in (X, Θ) space with $0 < x < 1$ and $-\pi/2 < \theta < \pi/2$. The values of X and Θ that

[8] First posed in 1733, its solution is given on pages 100–104 of Buffon, G., 1777: Essai d'arithmétique morale. *Histoire naturelle, générale et particulière, Supplément*, **4**, 46–123.

lead to the intersection with a parallel line is $0 < x < L\cos(\theta)$ where $-\pi/2 < \theta < \pi/2$. Consequently, from Equation 16.2.5,

$$p = \int_{-\pi/2}^{\pi/2} \int_0^{L\cos(\theta)} p_{X\Theta}(x,\theta)\, dx\, d\theta = \int_{-\pi/2}^{\pi/2} \int_0^{L\cos(\theta)} p_X(x) p_\Theta(\theta)\, dx\, d\theta \tag{16.7.33}$$

$$= \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \int_0^{L\cos(\theta)} dx\, d\theta = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} L\cos(\theta)\, d\theta = \frac{2L}{\pi}. \tag{16.7.34}$$

Consequently, given L, we can perform the tossing either physically or numerically, measure p, and compute the value of π. □

Convolution

It is often important to calculate the distribution of $X + Y$ from the distribution of X and Y when X and Y are independent. We shall derive the relationship for continuous random variables and then state the result for X and Y discrete.

Let X have a probability density function $p_X(x)$ and Y has the probability density $p_Y(y)$. Then the cumulative distribution function of $X + Y$ is

$$G_{X+Y}(a) = P(x + y \le a) = \int\int_{x+y\le a} p_X(x) p_Y(y)\, dx\, dy \tag{16.7.35}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} p_X(x) p_Y(y)\, dx\, dy = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{a-y} p_X(x)\, dx \right] p_Y(y)\, dy \tag{16.7.36}$$

$$= \int_{-\infty}^{\infty} F_X(a - y) p_Y(y)\, dx. \tag{16.7.37}$$

Therefore,

$$p_{X+Y}(a) = \frac{d}{da} \left[\int_{-\infty}^{\infty} F_X(a - y) p_Y(y)\, dy \right] = \int_{-\infty}^{\infty} p_X(a - y) p_Y(y)\, dy. \tag{16.7.38}$$

In the case when X and Y are discrete,

$$p_{X+Y}[a_k] = \sum_{i=-\infty}^{\infty} p_X[x_i] p_Y[a_k - x_i]. \tag{16.7.39}$$

Covariance

In Section 16.5 we introduced the concept of variance of a random variable X. There we showed that this quantity measures the dispersion, or spread, of the distribution of X about its expectation. What about the case of two jointly distributed random numbers?

Our first attempt might be to look at $\text{Var}(X)$ and $\text{Var}(Y)$. But this would simply display the dispersions of X and Y independently rather than jointly. Indeed, $\text{Var}(X)$ would give the spread along the x-direction while $\text{Var}(Y)$ would measure the dispersion along the y-direction.

Consider now $\text{Var}(aX+bY)$, the joint spread of X and Y along the $(ax+by)$-direction for two arbitrary real numbers a and b. Then

$$\begin{aligned}\text{Var}(aX+bY) &= E[(aX+bY)-E(aX+bY)]^2 &\quad (\mathbf{16.7.40})\\ &= E[(aX+bY)-E(aX)-E(bY)]^2 &\quad (\mathbf{16.7.41})\\ &= E\{a[X-E(X)]+b[Y-E(Y)]\}^2 &\quad (\mathbf{16.7.42})\\ &= E\{a^2[X-E(X)]^2+b^2[Y-E(Y)]^2+2ab[X-E(X)][Y-E(Y)]\} &\quad (\mathbf{16.7.43})\\ &= a^2\text{Var}(X)+b^2\text{Var}(Y)+2abE\{[X-E(X)][Y-E(Y)]\}. &\quad (\mathbf{16.7.44})\end{aligned}$$

Thus, the joint spread or dispersion of X and Y in any arbitrary direction $ax+by$ depends upon three parameters: $\text{Var}(X)$, $\text{Var}(Y)$, and $E\{[X-E(X)][Y-E(Y)]\}$. Because $\text{Var}(X)$ and $\text{Var}(Y)$ give the dispersion of X and Y separately, it is the quantity $E\{[X-E(X)][Y-E(Y)]\}$ that measures the joint spread of X and Y. This last quantity,

$$\text{Cov}(X,Y) = E\{[X-E(X)][Y-E(Y)]\}, \qquad (\mathbf{16.7.45})$$

is called the *covariance* and is usually denoted by $\text{Cov}(X,Y)$ because it determines how X and Y covary jointly. It only makes sense when we have two different random variables because in the case of a single random variable, $\text{Cov}(X,X)=\sigma_X^2=\text{Var}(X)$. Furthermore, $\text{Cov}(X,Y)\le\sigma_X\sigma_Y$. In summary,

$$\text{Var}(aX+bY) = a^2\,\text{Var}(X)+b^2\,\text{Var}(Y)+2ab\,\text{Cov}(X,Y). \qquad (\mathbf{16.7.46})$$

An alternative method for computing the covariance occurs if we recall that $\mu_X=E(X)$ and $\mu_Y=E(Y)$. Then

$$\begin{aligned}\text{Cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] = E(XY-\mu_XY-\mu_YX+\mu_X\mu_Y) &\quad (\mathbf{16.7.47})\\ &= E(XY)-\mu_XE(Y)-\mu_YE(X)+\mu_X\mu_Y &\quad (\mathbf{16.7.48})\\ &= E(XY)-\mu_X\mu_Y-\mu_Y\mu_X+\mu_X\mu_Y &\quad (\mathbf{16.7.49})\\ &= E(XY)-\mu_X\mu_Y = E(XY)-E(X)E(Y), &\quad (\mathbf{16.7.50})\end{aligned}$$

where

$$E(XY) = \begin{cases}\displaystyle\sum_{x_i\in A, y_j\in B} x_iy_j\,p_{XY}[x_i,y_j], & X \text{ discrete},\\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,p_{XY}(x,y)\,dx\,dy, & X \text{ continuous}.\end{cases} \qquad (\mathbf{16.7.51})$$

Therefore,

$$\text{Cov}(X,Y) = E(XY)-E(X)E(Y). \qquad (\mathbf{16.7.52})$$

It is left as a homework assignment to show that

$$\text{Cov}(aX+b,cY+d) = ac\,\text{Cov}(X,Y). \qquad (\mathbf{16.7.53})$$

In general, $\text{Cov}(X,Y)$ can be positive, negative, or zero. For it to be positive, X and Y decrease together or increase together. For a negative value, X would increase while Y decreases, or vice versa. If $\text{Cov}(X,Y)>0$, X and Y are *positively correlated*. If

$\mathrm{Cov}(X, Y) < 0$, X and Y are *negatively correlated.* Finally, if $\mathrm{Cov}(X, Y) = 0$, X and Y are *uncorrelated.*

• Example 16.7.6

The following table gives a discrete joint density function:

$p_{XY}[x_i, y_j]$	x_i: 0	1	2	$p_Y[y_j]$
y_j 0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
$p_X[x_i]$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	

Because

$$E(XY) = \sum_{i=0}^{2}\sum_{j=0}^{2} x_i\, y_j\, p_{XY}[x_i, y_j] = \frac{3}{14}, \tag{16.7.54}$$

$$\mu_X = E(X) = \sum_{i=0}^{2} x_i\, p_X[x_i] = \frac{3}{4}, \qquad \text{and} \qquad \mu_Y = E(Y) = \sum_{j=0}^{2} y_j\, p_Y[y_j] = \frac{1}{2}, \tag{16.7.55}$$

then

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{3}{14} - \frac{3}{4}\cdot\frac{1}{2} = -\frac{9}{56}. \tag{16.7.56}$$

Therefore, X and Y are *negatively* correlated. □

• Example 16.7.7

The random variables X and Y have the joint probability density function

$$p_{XY}(x, y) = \begin{cases} x + y, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{16.7.57}$$

Let us compute the covariance.

First, we must compute $p_X(x)$ and $p_Y(y)$. We find that

$$p_X(x) = \int_0^1 p_{XY}(x, y)\, dy = \int_0^1 (x + y)\, dy = x + \tfrac{1}{2} \tag{16.7.58}$$

for $0 < x < 1$, and

$$p_Y(y) = \int_0^1 p_{XY}(x, y)\, dx = \int_0^1 (x + y)\, dx = y + \tfrac{1}{2} \tag{16.7.59}$$

for $0 < y < 1$.

Because

$$E(XY) = \int_0^1 \int_0^1 xy(x+y)\,dx\,dy = \int_0^1 \left(y\frac{x^3}{3}\bigg|_0^1 + y^2\frac{x^2}{2}\bigg|_0^1 \right) dy \tag{16.7.60}$$

$$= \int_0^1 \left(\frac{y}{3} + \frac{y^2}{2}\right) dy = \frac{y^2}{6} + \frac{y^2}{6}\bigg|_0^1 = \frac{1}{3}, \tag{16.7.61}$$

$$\mu_X = E(X) = \int_0^1 x\,p_X(x)\,dx = \int_0^1 \left(x^2 + \frac{x}{2}\right) dx = \frac{7}{12}, \tag{16.7.62}$$

$$\mu_Y = E(Y) = \int_0^1 y\,p_Y(y)\,dy = \int_0^1 \left(y^2 + \frac{y}{2}\right) dy = \frac{7}{12}, \tag{16.7.63}$$

then

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}. \tag{16.7.64}$$

Therefore, X and Y are *negatively* correlated. □

Correlation

Although the covariance tells us how X and Y vary jointly, it depends upon the same units in which X and Y are measured. It is often better if we free ourselves of this nuisance, and we now introduce the concept of correlation.

Let X and Y be two random variables with $0 < \sigma_X^2 < \infty$ and $0 < \sigma_Y^2 < \infty$. The *correlation coefficient* $\rho(X,Y)$ between X and Y is given by

$$\rho(X,Y) = \text{Cov}\left[\frac{X - E(X)}{\sigma_X}, \frac{Y - E(Y)}{\sigma_Y}\right] = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}. \tag{16.7.65}$$

It is noteworthy that $|\rho(X,Y)| \le 1$.

Random Vectors

It is often useful to express our two random variables X and Y as a two-dimensional *random vector* $\mathbf{V} = (X\ \ Y)^T$. Then, the covariance can be written as 2×2 *covariance matrix*, given by

$$\begin{pmatrix} \text{cov}(X,X) & \text{cov}(X,Y) \\ \text{cov}(Y,X) & \text{cov}(Y,Y) \end{pmatrix}.$$

These considerations can be generalized into the n-dimensional *random vector* consisting of n random variables that are all associated with the same events.

• Example 16.7.7

Using MATLAB, let us create two random variables by invoking `X = randn(N,1)` and `Y = randn(N,2)`, where `N` is the sample size. If `N = 10`, we would find that using the MATLAB command `cov(X,Y)` would yield

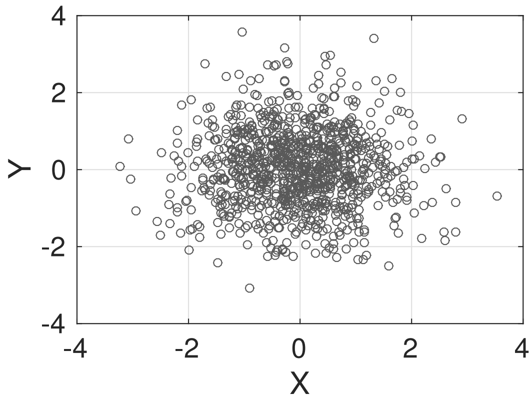


Figure 16.7.2: Scatter plot of points (X_i, Y_i) given by the random vector $\mathbf{V}$ in Example 16.7.7 when `N = 1000`.

```
>> ans =

          3.1325   0.9748
          0.9748   1.4862
```

.

(If you do this experiment, you will also obtain a symmetric matrix but with different elements.) On the other hand, if `N = 1000`, we find that `cov(X,Y)` equals

```
>> ans =
```

$$\begin{matrix} 0.9793 & -0.0100 \\ -0.0100 & 0.9927 \end{matrix}.$$

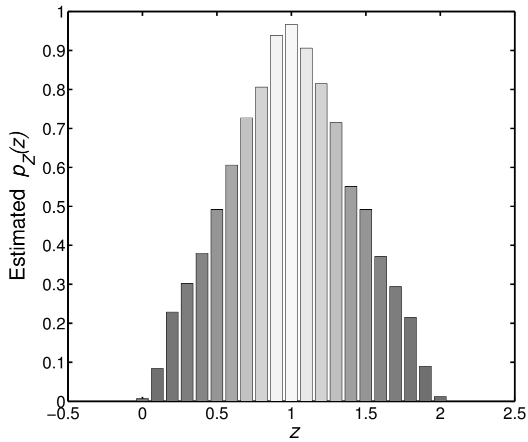
The interpretation of the covariance matrix is as follows: The variance (or spread) of data given by $\mathbf{X}$ and $\mathbf{Y}$ is (essentially) unity. The correlation between $\mathbf{X}$ and $\mathbf{Y}$ is (essentially) zero. These results are confirmed in Figure 16.7.2 where we have plotted X and Y as the data points (X_i, Y_i) when `N = 1000`. We can see the symmetric distribution of data points.

Problems

1. A search committee of 5 is selected from a science department that has 7 mathematics professors, 8 physics professors, and 5 chemistry professors. If X and Y denote the number of mathematics and physics professors, respectively, that are selected, calculate the joint probability function.

2. In an experiment of rolling a fair die twice, let Z denote a random variable that equals the sum of the results. What is $p_Z[z_i]$? Hint: Let X denote the result from the first toss and Y denote the result from the second toss. What you must find is $Z = X + Y$.

3. Show that $\text{Cov}(aX + b, cY + d) = ac\,\text{Cov}(X, Y)$.



Convolution Project

Project: Convolution

Consider two independent, uniformly distributed random variables (X, Y) that are summed to give $Z = X + Y$ with

$$p_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$p_Y(y) = \begin{cases} 1, & 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Show that

$$p_Z(z) = \begin{cases} z, & 0 < z \le 1, \\ 2 - z, & 1 < z \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

Then confirm your results using MATLAB's intrinsic function `rand` to generate $\{x_i\}$ and $\{y_j\}$ and computing $p_Z(z)$. You may want to review Example 11.5.1.

Further Readings

Beckmann, P., 1967: *Probability in Communication Engineering.* Harcourt, Brace & World, 511 pp. A presentation of probability as it applies to problems in communication engineering.

Ghahramani, S., 2000: *Fundamentals of Probability.* Prentice Hall, 511 pp. Nice introductory text on probability with a wealth of examples.

Hsu, H., 1997: *Probability, Random Variables, & Random Processes.* McGraw-Hill, 306 pp. Summary of results plus many worked problems.

Kay, S. M., 2006: *Intuitive Probability and Random Processes Using MATLAB*. Springer, 833 pp. A well-paced book designed for the electrical engineering crowd.

Ross, S. M., 2007: *Introduction to Probability Models.* Academic Press, 782 pp. An introductory undergraduate book in applied probability and stochastic processes.

Tuckwell, H. C., 1995: *Elementary Applications of Probability Theory.* Chapman & Hall, 292 pp. This book presents applications using probability theory, primarily from biology.

Chapter 17

Random Processes

In the previous chapter we introduced the concept of a random variable X. There X assumed various values of x according to a probability mass function $p_X[k]$ or probability density function $p_X(x)$. In this chapter we generalize the random variable so that it is also a function of time t. As before, the values of x assumed by the random variable $X(t)$ at a certain time is still unknown beforehand and unpredictable.

Our random, time-varying variable $X(t;\xi)$ is often used to describe a *stochastic* or *random process*. In that case, $X(t)$ is the *state* of the process at time t. The process can be either discrete or continuous in t.

A random process is not one function but a collection or family of functions, called *sample functions*, with some probability assigned to each. When we perform an experiment, we observe only one of these functions that is called a *realization* or *sample path* of the process. To observe more than a single function, we must repeat the experiment.

The *state space* of a random process is the set of *all* possible values that the random variable $X(t)$ can assume.

We can view random processes from many perspectives. First, it is a random function of time. This perspective is useful when we wish to relate an evolutionary physical phenomenon to its probabilistic model. Second, we can focus on its aspect as a random variable. This is useful in developing mathematical methods and tools to analyze random processes.

Another method for characterizing a random process examines its behavior as t and ξ vary or are kept constant. For example, if we allow t and ξ to vary, we obtain a family or *ensemble* of $X(t)$. If we allow t to vary while ξ is fixed, then $X(t)$ is simply a function of time and gives a sample function or *realization* for this particular random process. On the other hand, if we fix t and allow ξ to vary, $X(t)$ is a random variable equal to the state of the random process at time t. Finally, if we fix both t and ξ, then $X(t)$ is a number.

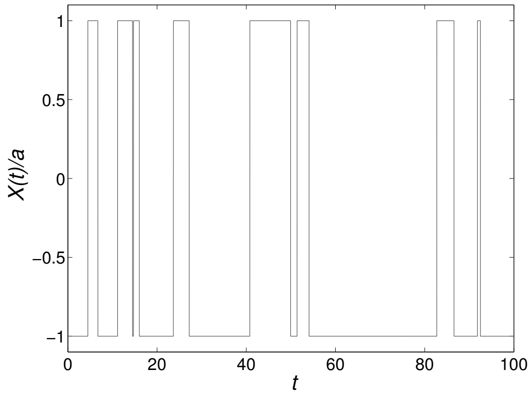


Figure 17.0.1: A realization of the random telegraph signal.

• Example 17.0.1

Consider a random process $X(t) = A$, where A is uniformly distributed in the interval $[0, 1]$. A plot of sample functions of $X(t)$ (a plot of $X(t)$ as a function of t) consists of horizontal straight lines that would cross the ordinate somewhere between 0 and 1. □

• Example 17.0.2

Consider the coin tossing experiment where the outcomes are either heads H or tails T. We can introduce the random process defined by

$$X(t; H) = \sin(t), \quad \text{and} \quad X(t; T) = \cos(t). \tag{17.0.1}$$

Note that the sample functions here are continuous functions of time. □

• Example 17.0.3: Random telegraph signal

Consider a signal that switches between $-a$ and $+a$ at random times. Suppose the process starts (at time $t = 0$) in the $-a$ state. It then remains in that state for a time interval T_1 at which point it switches to the state $X(t) = a$. The process remains in that state until $t = T_2$, then switches back to $X(t) = -a$. The switching time is given by a Poisson process, a random process that we discuss in Section 17.6. Figure 17.0.1 illustrates the random telegraph signal. □

Of all the possible random processes, a few are so useful in engineering and the physical sciences that they warrant special names. Some of them are:

• *Bernoulli process*

Imagine an electronics firm that produces electronic devices that either work (a success denoted by "S") or do not work (a failure or denoted "F"). We can model the production line as a series of independent, repeated events where p denotes the probability of producing a working device and $q = 1 - p$ is the probability of producing a faulty device. Thus, the production line can be modeled as random process, called a *Bernoulli process*, which has discrete states and parameter space.

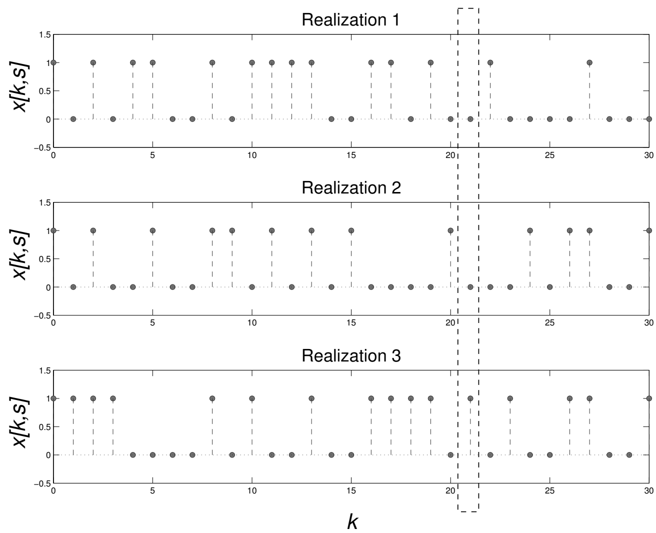


Figure 17.0.2: Three realization or sample functions of a Bernoulli random process with $p = 0.4$. The realization starts at $k = 0$ and continues forever. The dashed box highlights the values of the random variable $X[21, s]$.

If we denote each discrete trial by the integer k, a Bernoulli process generates successive outcomes at times $k = 0, 1, 2, \ldots$. Mathematically we can express this discrete random process by $X[k, s]$ where k denotes the time and s denotes the number of the realization or sample function. Furthermore, this random process maps the original experimental sample space $\{(F, F, S, \ldots), (S, F, F, \ldots), (F, F, F, \ldots), \ldots\}$ to the numerical sample space $\{(0, 0, 1, \ldots), (1, 0, 0, \ldots), (0, 0, 0, \ldots), \ldots\}$. Unlike the Bernoulli *trial* that we examined in the previous chapter, each simple event now becomes an infinite *sequence* of S's and F's.

Figure 17.0.2 illustrates three realizations or sample functions for a Bernoulli random variable when $p = 0.4$. In each realizations $s = 1, 2, \ldots$, the abscissa denotes time where each successive trial occurs at times $k = 0, 1, 2, \ldots$. When we fix the value of k, the quantity $X[k, s]$ is a random variable with a probability mass function of a Bernoulli random variable.

- *Markov process*

Communication systems transmit either the digits 0 or 1. Each transmitted digit often must pass through several stages. At each stage there is a chance that the digit that enters one stage will be changed by the time when it leaves.

A Markov process is a stochastic process that describes the probability that the digit will or will not be changed. It does this by computing the conditional distribution of any future state X_{n+1} by considering only the past states $X_0, X_1, \ldots, X_{n-1}$ and the present state X_n. In Section 17.4 we examine the simplest possible discrete Markov process, a Markov chain, when only the present and previous stages are involved. An example is the probabilistic description of birth and death, which is given in Section 17.5.

- *Poisson process*

The prediction of the total number of "events" that occur by time t is important to such diverse fields as telecommunications and banking. The most popular of these *counting processes* is the Poisson process. It occurs when:

1. The events occur "rarely."
2. The events occur in nonoverlapping intervals of time that are independent of each other.
3. The events occur at a constant rate λ.

In Section 17.6 we explore this random process.

- *Wiener process*

A Wiener process W_t is a random process that is continuous in time and possesses the following three properties:

1. $W_0 = 0$,
2. W_t is almost surely continuous, and
3. W_t has independent increments with a distribution $W_t - W_s \sim N(0, t-s)$ for $0 \le s \le t$.

As a result of these properties, we have that

1. the expectation is zero, $E(W_t) = 0$,
2. the variance is $E(W_t^2) - E^2(W_t) = t$, and
3. the covariance is $\text{cov}(W_s, W_t) = \min(s, t)$.

Norbert Wiener (1894–1964) developed this process to rigorously describe the physical phenomena of Brownian motion - the apparent random motion of particles suspended in a fluid. In a Wiener process the distances traveled in Brownian motion are distributed according to a Gaussian distribution and the path is continuous but consists entirely of sharp corners.

Project: Gambler's Ruin Problem

Pete and John decide to play a coin-tossing game. Pete agrees to pay John 10 cents whenever the coin yields a "head" and John agrees to pay Pete 10 cents whenever it is a "tail." Let S_n denote the amount that John earns in n tosses of a coin. This game is a stochastic process with discrete time (number of tosses). The state space is $\{0, \pm 10, \pm 20, \cdots\}$ cents. A realization occurs each time that they play a new game.

Step 1: Create a MATLAB code to compute a realization of S_n. Plot several realizations (sample functions) of this random process. See Figure 17.0.3.

Step 2: Suppose Pete has 10 dimes. Therefore, there is a chance he will run out of dimes at some $n = N$. Modify your MATLAB code to construct a probability density function that gives the probability Pete will run out of money at time $n = N$. See Figure 17.0.3. This is an example of the famous *gambler's ruin problem*.

17.1 FUNDAMENTAL CONCEPTS

In Section 16.5 we introduced the concepts of mean (or expectation) and variance as it applies to discrete and continuous random variables. These parameters provide useful characterizations of a probability mass function or probability density function. Similar considerations hold in the case of random processes and we introduce them here.

Mean and variance

We define the mean of the random process $X(t)$ as the expected value of the process - that is, the expected value of the random variable defined by $X(t)$ for a fixed instant of

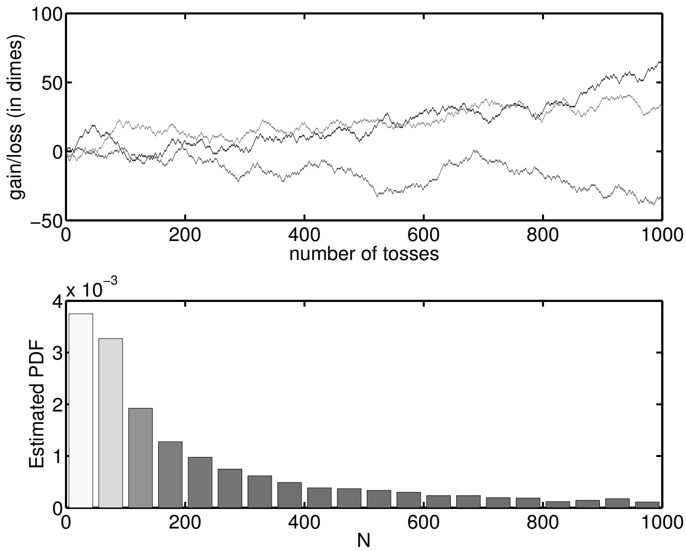


Figure 17.0.3: (a) Top frame: John's gains or losses as the result of the three different coin tossing games. (b) The probability density function for John's winning 10 dimes as a function of the number of tosses that are necessary to win 10 dimes.

time. Note that when we take the expectation, we hold the time as a nonrandom parameter and average only over the random quantities. We denote this mean of the random process by $\mu_X(t)$, since, in general, it may depend on time. The definition of the mean is just the expectation of $X(t)$:

$$\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x\, p_{X(t)}(t;x)\,dx. \tag{17.1.1}$$

In a similar vein, we can generalize the concept of variance so that it applies to random processes. Here variance also becomes a time-dependent function defined by

$$\sigma_X^2(t) = \text{Var}[X(t)] = E\left\{[X(t) - \mu_X(t)]^2\right\}. \tag{17.1.2}$$

• Example 17.1.1: Random linear trajectories

Consider the random process defined by

$$X(t) = A + Bt, \tag{17.1.3}$$

where A and B are uncorrelated random variables with means μ_A and μ_B. Let us find the mean of this random process.

From the linearity property of expectation, we have that

$$\mu_X(t) = E[X(t)] = E(A + Bt) = E(A) + E(B)t = \mu_A + \mu_B t. \tag{17.1.4}$$

□

• **Example 17.1.2: Random sinusoidal signal**

A random sinusoidal signal is one governed by $X(t) = A\cos(\omega_0 t + \Theta)$, where A and Θ are *independent* random variables, A has a mean μ_A and variance σ_A^2, and Θ has the probability density function $p_\Theta(x)$ that is nonzero only over the interval $(0, 2\pi)$.

The mean of $X(t)$ is given by

$$\mu_X(t) = E[X(t)] = E[A\cos(\omega_0 t + \Theta)] = E[A]E[\cos(\omega_0 t + \Theta)]. \qquad \textbf{(17.1.5)}$$

We have used the property that the expectation of two independent random variables equals the product of the expectation of each of the random variables. Simplifying Equation 17.1.5,

$$\mu_X(t) = \mu_A \int_0^{2\pi} \cos(\omega_0 t + x) p_\Theta(x)\, dx. \qquad \textbf{(17.1.6)}$$

A common assumption is that $p_\Theta(x)$ is uniformly distributed in the interval $(0, 2\pi)$, namely

$$p_\Theta(x) = \frac{1}{2\pi}, \qquad 0 < x < 2\pi. \qquad \textbf{(17.1.7)}$$

Substituting Equation 17.1.7 into Equation 17.1.6, we find that

$$\mu_X(t) = \frac{\mu_A}{2\pi} \int_0^{2\pi} \cos(\omega_0 t + x)\, dx = 0. \qquad \textbf{(17.1.8)}$$

□

• **Example 17.1.3: Wiener random process or Brownian motion**

A Wiener (random) process is defined by

$$X(t) = \int_0^t U(\xi)\, d\xi, \qquad t \ge 0, \qquad \textbf{(17.1.9)}$$

where $U(t)$ denotes white Gaussian noise. It is often used to model Brownian motion. To find its mean, we have that

$$E[X(t)] = E\left[\int_0^t U(\xi)\, d\xi\right] = \int_0^t E[U(\xi)]\, d\xi = 0, \qquad \textbf{(17.1.10)}$$

because the mean of white Gaussian noise equals zero. □

Autocorrelation function

When a random process is examined at two time instants $t = t_1$ and $t = t_2$, we obtain two random variables $X(t_1)$ and $X(t_2)$. A useful relationship between these two random variables is found by computing their correlation as a function of time instants t_1 and t_2. Because it is a correlation between the values of the same process sampled at two different instants of time, we shall call it the *autocorrelation function* of the process $X(t)$ and denote it by $R_X(t_1, t_2)$. It is defined in the usual way for expectations by

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)]. \qquad \textbf{(17.1.11)}$$

Just as in the two random variables case, we can define the covariance and correlation coefficient, but here the name is slightly different. We define the *autocovariance function* as

$$C_X(t_1, t_2) = E\{[X(t_1) - \mu_X(t_1)][X(t_2) - \mu_X(t_2)]\} \qquad \textbf{(17.1.12)}$$
$$= R_X(t_1, t_2) - \mu_X(t_1)\mu_X(t_2). \qquad \textbf{(17.1.13)}$$

Note that the variance of the process and its average power (the names used for the average of $[X(t)-\mu_X(t)]^2$ and $[X(t)]^2$, respectively) can be directly obtained for the autocorrelation and the autocovariance functions, by simply using the same time instants for both t_1 and t_2:

$$E\{[X(t)]^2\} = R_X(t, t), \qquad \textbf{(17.1.14)}$$

and

$$\sigma_X^2(t) = E\{[X(t) - \mu_X(t)]^2\} = C_X(t, t) = R_X(t, t) - \mu_X^2(t). \qquad \textbf{(17.1.15)}$$

Therefore, the average power, Equation 17.1.14, and the variance, Equation 17.1.15, of the process follows directly from the definition of the autocorrelation and autocovariance functions.

• Example 17.1.4: Random linear trajectories

Let us continue Example 17.1.1 and find the autocorrelation of a random linear trajectory given by $X(t) = A + Bt$. From the definition of the autocorrelation,

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = E\{[A + Bt_1][A + Bt_2]\} \qquad \textbf{(17.1.16)}$$
$$= E(A^2) + E(AB)(t_1 + t_2) + E(B^2)t_1t_2 \qquad \textbf{(17.1.17)}$$
$$= (\sigma_A^2 + \mu_A^2) + \mu_A\mu_B(t_1 + t_2) + (\sigma_B^2 + \mu_B^2)t_1t_2, \qquad \textbf{(17.1.18)}$$

where σ_A^2 and σ_B^2 are the variances of the random variables A and B. We can easily find the autocovariance by

$$C_X(t_1, t_2) = R_X(t_1, t_2) - \mu_X(t_1)\mu_X(t_2) = \sigma_A^2 + \sigma_B^2 t_1t_2. \qquad \textbf{(17.1.19)}$$

□

• Example 17.1.5: Random sinusoidal signal

We continue to examine the random sinusoidal signal given by $X(t) = A\cos(\omega_0 t + \Theta)$. The autocorrelation function is

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = E[A\cos(\omega_0 t_1 + \Theta)A\cos(\omega_0 t_2 + \Theta)] \qquad \textbf{(17.1.20)}$$
$$= \tfrac{1}{2}E(A^2)E[\cos(\omega_0 t_2 - \omega_0 t_1) + \cos(\omega_0 t_2 + \omega_0 t_1 + 2\Theta)] \qquad \textbf{(17.1.21)}$$
$$= \tfrac{1}{2}(\sigma_A^2 + \mu_A^2)\left\{\cos[\omega_0(t_2 - t_1)] + \int_0^{2\pi} \cos[\omega_0(t_2 + t_1) + 2x]p_\Theta(x)\,dx\right\}. \textbf{(17.1.22)}$$

In our derivation we used (1) the property that the expectation of A^2 equals the sum of the variance and the square of the mean, and (2) the first term involving the cosine is *not* random because it is a function of only the time instants and the frequency. From Equation 17.1.22 we see that autocorrelation function may depend on both time instants if the probability density function of the phase angle is arbitrary. On the other hand, if $p_\Theta(x)$ is uniformly distributed, then the last term in Equation 17.1.22 vanishes because integrating the cosine function over the interval of one period is zero. In this case we can write the autocorrelation function as a function of only the time difference. The process also becomes wide-sense stationary with

$$R_X(\tau) = E[X(t)X(t+\tau)] = \tfrac{1}{2}(\sigma_A^2 + \mu_A^2)\cos(\omega_0\tau). \tag{17.1.23}$$

□

Wide-sense stationary processes

A process is strictly stationary if its distribution and density functions do not depend on the absolute values of the time instants t_1 and t_2 but only on the difference of the time instants, $|t_1 - t_2|$. However, this is a very rigorous condition. If we are concerned only with the mean and autocorrelation function, then we can soften our definition of a stationary process to a limited form, and we call such processes wide-sense stationary processes. A *wide-sense stationary process* has a constant mean, and its autocorrelation function depends only on the time difference:

$$\mu_X(t) = E[X(t)] = \mu_X, \tag{17.1.24}$$

and

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = R_X(t_2 - t_1). \tag{17.1.25}$$

Because time does not appear in the mean, we simply write it as a constant mean value μ_X. Similarly, because the autocorrelation function is a function only of the time difference, we can write it as a function of a single variable, the time difference τ:

$$R_X(\tau) = E[X(t)X(t+\tau)]. \tag{17.1.26}$$

We can obtain similar expressions for the autocovariance function, which in this case depends only on the time difference as well:

$$C_X(\tau) = E\{[X(t) - \mu_X][X(t+\tau) - \mu_X]\} = R_X(\tau) - \mu_X^2. \tag{17.1.27}$$

Finally, the average power and variance for a wide-sense stationary process are

$$E\{[X(t)]^2\} = R_X(0), \tag{17.1.28}$$

and

$$\sigma_X^2 = C_X(0) = R_X(0) - \mu_X^2, \tag{17.1.29}$$

respectively. Therefore, a wide-sense stationary process has a constant average power and constant variance.

Problems

1. Find $\mu_X(t)$ and $\sigma_X^2(t)$ for the random process given by $X(t) = A\cos(\omega t)$, where ω is a constant and A is a random variable with the Gaussian (or normal) probability density function

$$p_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

2. Consider a sine-wave random process $X(t) = A\cos(\omega t + \Theta)$, where A and ω are constants with $A > 0$. The phase function Θ is a random, uniform variable on the interval $[-\pi, \pi]$. Find the mean and autocorrelation for this random function.

3. Consider a countably infinite sequence $\{X_n, n = 0, 1, 2, 3, \ldots\}$ of a random variable defined by

$$X_n = \begin{cases} 1, & \text{for success in the } n\text{th trial}, \\ 0, & \text{for failure in the } n\text{th trial}, \end{cases}$$

with the probabilities $P(X_n = 0) = 1 - p$ and $P(X_n = 1) = p$. Thus, X_n is a Bernoulli process. For this process, $E(X_n) = p$ and $\text{Var}(X_n) = p(1-p)$. Show that the autocorrelation is

$$R_X(t_1, t_2) = \begin{cases} p, & t_1 = t_2, \\ p^2, & t_1 \neq t_2; \end{cases}$$

and the autocovariance is

$$C_X(t_1, t_2) = \begin{cases} p(1-p), & t_1 = t_2, \\ 0, & t_1 \neq t_2. \end{cases}$$

Project: Computing the Autocorrelation Function

In most instances you must compute the autocorrelation function numerically. The purpose of this project is to explore this computation using the random telegraph signal. The exact solution is given by Equation 17.2.24. You will compute the autocorrelation two ways:

Step 1: Using Example 17.6.1, create MATLAB code that generates 500 realizations of the random telegraph signal.

Step 2: Choosing an arbitrary time t_S, compute $X_k(t_S)X_k(t_S + \tau)$ for $0 \leq 0 \leq \tau_{max}$ and $k = 1, 2, 3, \ldots, 500$. Then find the average value of $X_k(t_S)X_k(t_S + \tau)$. Plot $R_X(\tau)$ as a function of τ and include the exact answer for comparison. Does it matter how many sample functions you use?

Step 3: Now introduce a number of times $t_m = m\Delta t$, where $m = 0, 1, 2, \ldots, M$. Using only a *single realization* $k = K$ of your choice, compute $X_K(m\Delta t) \times X_K(m\Delta t + \tau)$. Then find the average value of $X_K(m\Delta t)X_K(m\Delta t + \tau)$ and plot this result as a function of τ. On the same plot, include the exact solution. Does the value of Δt matter? See Figure 17.1.1

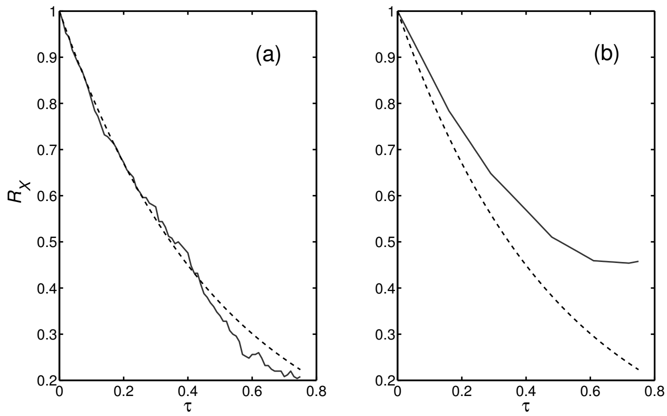


Figure 17.1.1: The autocorrelation function $R_X(\tau)$ for the random telegraph signal as a function of τ when $\lambda = 2$. The dashed line gives the exact solution. In frame (a) $X_k(t_S)X_k(t_S+\tau)$ has been averaged over 500 realizations when $t_S = 2$. In frame (b) $X_{200}(m\Delta t)X_{200}(m\Delta t + \tau)$ has been averaged with $M = 1200$ and $\Delta t = 0.01$.

17.2 POWER SPECTRUM

In earlier chapters we provided two alternative descriptions of signals, either in the time domain, which provides information on the shape of the waveform, or in the frequency domain, which provides information on the frequency content. Because random signals do not behave in any predictable fashion nor are they represented by a single function, it is unlikely that we can define the spectrum of a random signal by taking its Fourier transform. On the other hand, the autocorrelation of random signals describes in some sense whether the signal changes rapidly or slowly. In this section we explain and illustrate the concept of power spectrum of random signals.

For a wide-sense stationary random signal $X(t)$ with autocorrelation function $R_X(\tau)$, the *power spectrum* $S_X(\omega)$ of the random signal is the Fourier transform of the autocorrelation function:

$$S_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) e^{-i\omega\tau}\, d\tau. \qquad \textbf{(17.2.1)}$$

Consequently, the autocorrelation can be obtained from inverse Fourier transform of the power spectrum, or

$$R_X(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_X(\omega) e^{i\omega\tau}\, d\omega. \qquad \textbf{(17.2.2)}$$

As with any Fourier transform, it enjoys certain properties. They are:

1. The power spectrum is real and even: $S_X(-\omega) = S_X(\omega)$ and $S_X^*(\omega) = S_X(\omega)$, where $S_X^*(\omega)$ denotes the complex conjugate value of $S_X(\omega)$.

2. The power spectrum is nonnegative: $S_X(\omega) \geq 0$.

3. The average power of the random signal is equal to the integral of the power spectrum:

$$E\{[X(t)]^2\} = R_X(0) = \frac{1}{\pi} \int_0^{\infty} S_X(\omega)\, d\omega. \qquad \textbf{(17.2.3)}$$

4. If the random signal has nonzero mean μ_X, then its power spectrum contains an impulse at zero frequency of magnitude $2\pi\mu_X^2$.
5. The Fourier transform of the autocovariance function of the random process is itself also a power spectrum and usually does not contain an impulse component in zero frequency.

Consider the following examples of the power spectrum:

• **Example 17.2.1: Random sinusoidal signal**

The sinusoidal signal is defined by

$$X(t) = A\cos(\omega_0 t + \Theta), \tag{17.2.4}$$

where the phase is uniformly distributed in the interval $[0, 2\pi]$. If the amplitude A has a mean of zero and a variance of σ^2, then the autocorrelation function is

$$R_X(\tau) = \tfrac{1}{2}\sigma^2\cos(\omega_0\tau) = R_X(0)\cos(\omega_0\tau). \tag{17.2.5}$$

The power spectrum of this signal is then

$$S_X(\omega) = \int_{-\infty}^{\infty} R_X(0)\cos(\omega_0\tau)e^{-i\omega\tau}\,d\tau = R_X(0)\pi\left[\delta(\omega-\omega_0) + \delta(\omega+\omega_0)\right]. \tag{17.2.6}$$

Because this signal contains only one frequency ω_0, its power spectrum is just two impulses, one at ω_0 and one at $-\omega_0$. Since the negative frequency appears due only to the even property of the power spectrum, it is clear that all power is concentrated at the frequency of the sinusoidal signal. While this is a very simple example, it does illustrate that the power spectrum indeed represents the way the power in the random signal is distributed among the various frequencies. We shall see later that if we also use linear systems in order to amplify or attenuate certain frequencies, the results mirror what we expect in the deterministic case. □

• **Example 17.2.2: Modulated signal**

Let us now examine a sinusoidal signal modulated by another random signal that contains low frequencies. This random process is described by

$$Y(t) = X(t)\cos(\omega_0 t + \Theta), \tag{17.2.7}$$

where the phase angle in Equation 17.2.7 is a random variable that is uniformly distributed in the interval $[0, 2\pi]$ and is independent of $X(t)$. Then the autocorrelation function of $Y(t)$ is given by

$$R_Y(\tau) = E[Y(t)Y(t+\tau)] = E\{X(t)\cos(\omega_0 t+\Theta)X(t+\tau)\cos[\omega_0(t+\tau)+\Theta]\} \tag{17.2.8}$$
$$= E[X(t)X(t+\tau)]E\{\cos(\omega_0 t+\Theta)\cos[\omega_0(t+\tau)+\Theta]\} = \tfrac{1}{2}R_X(\tau)\cos(\omega_0 t). \tag{17.2.9}$$

Let us take $R_X(\tau) = R_X(0)e^{-2\lambda|\tau|}$, the autocorrelation function for a random telegraph signal (see Equation 17.2.22). In this case,

$$R_Y(\tau) = \tfrac{1}{2}R_X(0)e^{-2\lambda|\tau|}\cos(\omega_0 t). \tag{17.2.10}$$

Turning to the power spectrum, the definition gives

$$S_Y(\omega) = \int_{-\infty}^{\infty} \tfrac{1}{2} R_X(\tau)\cos(\omega_0 t)e^{-i\omega\tau}\,d\tau \tag{17.2.11}$$

$$= \tfrac{1}{4}\int_{-\infty}^{\infty} R_X(\tau)\left(e^{i\omega_0\tau} + e^{-i\omega_0\tau}\right)e^{-i\omega\tau}\,d\tau \tag{17.2.12}$$

$$= \tfrac{1}{4}\left[S_X(\omega-\omega_0) + S_X(\omega+\omega_0)\right]. \tag{17.2.13}$$

Thus, the resulting power spectrum is shifted to the modulating frequency ω_0 and its negative value, with peak values located at both $\omega = \omega_0$ and $\omega = -\omega_0$.

□

• Example 17.2.3: White noise

There are instances when we want to approximate random signals where the autocorrelation function is very narrow and very large about $\tau = 0$. In those cases we construct an idealization of the autocorrelation function by using the impulse or delta function $\delta(\tau)$.

In the present case when $R_X(\tau) = C\,\delta(\tau)$, the power spectrum is

$$S_X(\omega) = \int_{-\infty}^{\infty} C\,\delta(\tau)e^{-i\omega\tau}\,d\tau = C. \tag{17.2.14}$$

Thus, the power spectrum here is a flat spectrum whose value is equal to C. Because the power spectrum is a flat for all frequencies, it is often called “white noise” since it contains all frequencies with equal weight.

An alternative derivation involves the random telegraph that we introduced in Example 17.0.3. As the switching rate becomes large and the rate λ approaches infinity, its amplitude increases as $\sqrt{\lambda}$. Because $R_X(0)$ increases linearly with λ, the autocorrelation function becomes

$$R_X(\tau) = C\lambda\exp(-2\lambda|\tau|). \tag{17.2.15}$$

The resulting power spectrum equals

$$S_X(\omega) = \lim_{\lambda\to\infty} \frac{4C\lambda^2}{\omega^2 + 4\lambda^2} = \lim_{\lambda\to\infty} \frac{C}{1 + [\omega/(2\lambda)]^2} = C. \tag{17.2.16}$$

The power spectrum is again flat for all frequencies.

The autocorrelation for white noise is an idealization because it has infinite average power. Obviously no real signal has infinite power since in practice the power spectrum decays eventually. Nevertheless, white noise is still quite useful because the decay usually occurs at such high frequencies that we can tolerate the errors of introducing a flat spectrum. □

• Example 17.2.4: Random telegraph signal

In Example 17.0.3 we introduced the random telegraph signal: $X(t)$ equals either $+h$ or $-h$, changing its value from one to the other in Poisson-distributed moments of time. The probability of n changes in a time interval τ is

$$P_\tau(n) = \frac{(\lambda\tau)^n}{n!}e^{-\lambda\tau}, \tag{17.2.17}$$

where λ denotes the average frequency of changes.

To compute the power spectrum we must first compute the correlation function via the product $X(t)X(t+\tau)$. This product equals h^2 or $-h^2$, depending on whether $X(t) = X(t+\tau)$ or $X(t) = -X(t+\tau)$, respectively. These latter relationships depend on the number of changes during the time interval. Now,

$$P[X(t) = X(t+\tau)] = P_\tau(n \text{ even}) = e^{-\lambda\tau} \sum_{n=1}^{\infty} \frac{(\lambda\tau)^{2n}}{(2n)!} = e^{-\lambda\tau} \cosh(\lambda\tau), \qquad \textbf{(17.2.18)}$$

and

$$P[X(t) = -X(t+\tau)] = P_\tau(n \text{ odd}) = e^{-\lambda\tau} \sum_{n=1}^{\infty} \frac{(\lambda\tau)^{2n+1}}{(2n+1)!} = e^{-\lambda\tau} \sinh(\lambda\tau). \qquad \textbf{(17.2.19)}$$

Therefore,

$$E[X(t)X(t+\tau)] = h^2 P_\tau(n \text{ even}) - h^2 P_\tau(n \text{ odd}) \qquad \textbf{(17.2.20)}$$
$$= h^2 e^{-\lambda\tau}[\cosh(\lambda\tau) - \sinh(\lambda\tau)] \qquad \textbf{(17.2.21)}$$
$$= h^2 e^{-2\lambda|\tau|}. \qquad \textbf{(17.2.22)}$$

We have introduced the absolute value sign in Equation 17.2.24 because our derivation was based on $t_2 > t_1$ and the absolute value sign takes care of the case $t_2 < t_1$.

Using Problem 1, we have that

$$S_X(\omega) = 2h^2 \int_0^\infty e^{-2\lambda\tau} \cos(\lambda\tau)\, d\tau = \frac{4h^2\lambda}{\omega^2 + 4\lambda^2}. \qquad \textbf{(17.2.23)}$$

Problems

1. Show that

$$S_X(\omega) = 2\int_0^\infty R_X(\tau)\cos(\omega\tau)\, d\tau.$$

17.3 TWO-STATE MARKOV CHAINS

A Markov chain is a probabilistic model in which the outcomes of successive trials depend only on its immediate predecessors. The mathematical description of a Markov chain involves the concepts of *states* and *state transition*. If $X_n = i$, then we have a process with state i and time n. Given a process in state i, there is a fixed probability P_{ij} that state i will transition into state j. In this section we focus on the situation of just two states.

Imagine that you want to predict the chance of rainfall tomorrow.[1] From close observation you note that the chance of rain tomorrow depends only on whether it is raining today and *not* on past weather conditions. From your observations you find that if it rains today, then it will rain tomorrow with probability α, and if it does not rain today, then the chance it will rain tomorrow is β. Assuming that these probabilities of changes are stationary (unchanging), you would like to answer the following questions:

---

[1] See, for example, Gabriel, K. R., and J. Neumann, 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quart. J. R. Met. Soc.*, **88**, 90–95.

1. Given that is raining (or not raining), what are the chances of it raining in eight days?
2. Suppose the day is rainy (or dry), how long will the current weather remain before it changes for the first time?
3. Suppose it begins to rain during the week, how long does it take before it stops?

If the weather observation takes place at noon, we have a discrete parameter process; the two possible states of the process are rain and no rain. Let these be denoted by 0 for no rain and 1 for rain. The four possible transitions are $(0 \to 0)$, $(0 \to 1)$, $(1 \to 0)$, and $(1 \to 1)$. Let X_n be the state of the process at the nth time point. We have $X_n = 0, 1$. Clearly, $\{X_n, n = 0, 1, 2, \ldots\}$ is a two-state Markov chain. Therefore questions about precipitation can be answered if all the properties of the two-state Markov chains are known. Let

$$P_{i,j}^{(m,n)} = P(X_n = j | X_m = i), \qquad i, j = 0, 1; \quad m \le n. \tag{17.3.1}$$

$P_{i,j}^{(m,n)}$ denotes the probability that the state of the process at the nth time point is j given that it was at state i at the mth time point. Furthermore, if this probability is larger for $i = j$ than when $i \neq j$, the system prefers to stay or *persist* in whatever state it is. When $n = m + 1$, we have that

$$P_{i,j}^{(m,m+1)} = P(X_{m+1} = j | X_m = i). \tag{17.3.2}$$

This is known as the one-step transition probability, given that the process is at i at time m.

There are two possibilities: either $P_{ij}^{(m,m+1)}$ depends on m or $P_{ij}^{(m,m+1)}$ is independent of m, where m is the initial value of the time parameter. Our precipitation model is an example of a second type of process in which the one-step transition probabilities do not change with time. Such processes are known as *time homogeneous*. Presently we shall restrict ourselves only to these processes. Consequently, without loss of generality we can use the following notation for the probabilities:

$$P_{ij} = P(X_{m+1} = j | X_m = i) \qquad \text{for all } m, \tag{17.3.3}$$

and

$$P_{ij}^{(n)} = P(X_{m+n} = j | X_m = i) \qquad \text{for all } m. \tag{17.3.4}$$

Chapman-Kolmogorov equation

The Chapman[2]-Kolmogorov[3] equations provide a mechanism for computing the transition probabilities after n steps. The n-step transition probabilities $P_{ij}^{(n)}$ denote the probability that a process in state i will be in state j after n transitions, or

$$P_{ij}^{(n)} = P[X_{n+k} = j | X_k = i], \qquad n \ge 0, \quad i, j \ge 0. \tag{17.3.5}$$

[2] Chapman, S., 1928: On the Brownian displacements and thermal diffusion of grains suspended via non-uniform fluid. *Proc. R. Soc. London, Ser. A*, **119**, 34–54.

[3] Kolmogorov, A. N., 1931: Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.*, **104**, 415–458.

Therefore, $P_{ij}^{(1)} = P_{ij}$. The Chapman-Kolmogorov equations give a method for computing these n-step transition probabilities via

$$P_{ij}^{(n+m)} = \sum_{k=0}^{\infty} P_{ik}^{(n)} P_{kj}^{(m)}, \qquad n, m \geq 0, \tag{17.3.6}$$

for all i and j. Here $P_{ik}^{(n)} P_{kj}^{(m)}$ represents the probability that the ith starting process will go to state j in $n+m$ transitions via a path that takes it into state k at the nth transition. Equation 17.3.6 follows from

$$P_{ij}^{(n+m)} = P[X_{n+m} = j | X_0 = i] = \sum_{k=0}^{\infty} P[X_{n+m} = j, X_n = k | X_0 = i] \tag{17.3.7}$$

$$= \sum_{k=0}^{\infty} P[X_{n+m} = j | X_n = k, X_0 = i] P[X_n = k | X_0 = i] = \sum_{k=0}^{\infty} P_{ik}^{(n)} P_{kj}^{(m)}. \tag{17.3.8}$$

Transmission probability matrix

Returning to the task at hand, we have that

$$P^{(2)} = P^{(1+1)} = P \cdot P = P^2, \tag{17.3.9}$$

and by induction

$$P^{(n)} = P^{(n-1+1)} = P^{(n-1)} \cdot P = P^n, \tag{17.3.10}$$

where $P^{(n)}$ denotes the transition matrix after n steps.

From our derivation, we see that: (1) The one-step transition probability matrix completely defines the time-homogeneous two-state Markov chain. (2) All transition probability matrices show the important property that the elements in any of their rows add up to one. This follows from the fact that the elements of a row represent the probabilities of mutually exclusive and exhaustive events on a sample space.

For two-state Markov processes, this means that

$$P_{00}^{(n)} + P_{01}^{(n)} = 1, \quad \text{and} \quad P_{10}^{(n)} + P_{11}^{(n)} = 1. \tag{17.3.11}$$

Furthermore, with the one-step transmission probability matrix:

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}, \quad 0 \leq a, b \leq 1, \quad |1-a-b| < 1, \tag{17.3.12}$$

then the n-step transmission probability matrix is

$$\begin{pmatrix} P_{00}^{(n)} & P_{01}^{(n)} \\ P_{10}^{(n)} & P_{11}^{(n)} \end{pmatrix} = \begin{pmatrix} \frac{b}{a+b} + a\frac{(1-a-b)^n}{a+b} & \frac{a}{a+b} - a\frac{(1-a-b)^n}{a+b} \\ \frac{b}{a+b} - b\frac{(1-a-b)^n}{a+b} & \frac{a}{a+b} + b\frac{(1-a-b)^n}{a+b} \end{pmatrix}. \tag{17.3.13}$$

This follows from the Chapman-Kolmogorov equation that

$$P_{00}^{(1)} = 1 - a, \tag{17.3.14}$$

Table 17.3.1: The Probability of Rain on the nth Day.

n	P_{00}	P_{10}	P_{01}	P_{11}
1	0.7000	0.2000	0.3000	0.8000
2	0.5500	0.3000	0.4500	0.7000
3	0.4750	0.3500	0.5250	0.6500
4	0.4375	0.3750	0.5625	0.6250
5	0.4187	0.3875	0.5813	0.6125
6	0.4094	0.3938	0.5906	0.6063
7	0.4047	0.3969	0.5953	0.6031
8	0.4023	0.3984	0.5977	0.6016
9	0.4012	0.3992	0.5988	0.6008
10	0.4006	0.3996	0.5994	0.6004
∞	0.4000	0.4000	0.6000	0.6000

and

$$P_{00}^{(n)} = (1-a)P_{00}^{(n-1)} + bP_{01}^{(n-1)}, \qquad n > 1, \tag{17.3.15}$$

$$= b + (1-a-b)P_{00}^{(n-1)}, \tag{17.3.16}$$

since $P_{01}^{(n)} = 1 - P_{00}^{(n)}$. Solving these equations recursively for $n = 1, 2, 3, \ldots$ and simplifying, we obtain Equation 17.3.13 as long as both a and b do not equal zero.

• Example 17.3.1

Consider a precipitation model where the chance for rain depends only on whether it rained yesterday. If we denote the occurrence of rain by state 0 and state 1 denotes no rain, then observations might give you a transition probability that looks like:

$$P = \begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix}. \tag{17.3.17}$$

Given that the atmosphere starts today with any one of these states, the probability of finding that it is raining on the nth day is given by P^n. Table 17.3.1 illustrates the results as a function of n. Thus, regardless of whether it rains today or not, in ten days the chance for rain is 0.4 while the chance for no rain is 0.6. □

Limiting behavior

As Table 17.3.1 suggests, as our Markov chain evolves, it reaches some steady-state. Let us explore this limit of $n \to \infty$ because it often provides a simple and insightful representation of a Markov process.

For large values of n it is possible to show that the limiting probability distribution of states is independent of the initial value. In particular, for $|1 - a - b| < 1$, we have that

$$\lim_{n\to\infty} P^{(n)} = \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}. \tag{17.3.18}$$

This follows from $\lim_{n\to\infty}(1-a-b)^n \to 0$ since $|1-a-b| < 1$. From Equation 17.3.13 the second term in each of the elements of the matrix tends to zero as $n \to \infty$.

Let us denote these limiting probabilities by $\pi_j = \lim_{n\to\infty} P_{ij}^{(n)}$. Then, from Equation 17.3.18,

$$\pi_{00} = \pi_{10} = \frac{b}{a+b} = \pi_0, \quad \text{and} \quad \pi_{01} = \pi_{11} = \frac{a}{a+b} = \pi_1, \tag{17.3.19}$$

and these limiting distributions are independent of the initial state.

Number of visits to a certain state

When a random process visits several states we would like to know the number of visits to a certain state. Let $N_{ij}^{(n)}$ denote the number of visits the two-state Markov chain $\{X_n\}$ makes to state j, starting initially at state i, in n time periods. If $\mu_{ij}^{(n)}$ denotes the expected number of visits that the process makes to state j in n steps after it originally started at state i, and the transition probability matrix P of the two-state Markov chain is

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \tag{17.3.20}$$

with $|1-a-b| < 1$, then

$$||\mu_{ij}^{(n)}|| = \begin{pmatrix} \frac{nb}{a+b} + \frac{a(1-a-b)[1-(1-a-b)^n]}{(a+b)^2} & \frac{na}{a+b} - \frac{a(1-a-b)[1-(1-a-b)^n]}{(a+b)^2} \\ \frac{nb}{a+b} - \frac{b(1-a-b)[1-(1-a-b)^n]}{(a+b)^2} & \frac{na}{a+b} + \frac{b(1-a-b)[1-(1-a-b)^n]}{(a+b)^2} \end{pmatrix} \tag{17.3.21}$$

To prove Equation 17.3.21, we introduce a random variable $Y_{ij}^{(k)}$, where

$$Y_{ij}^{(n)} = \begin{cases} 1, & \text{if } X_k = j \text{ and } X_0 = i, \\ 0, & \text{otherwise,} \end{cases} \tag{17.3.22}$$

for $i, j = 0, 1$. This random variable $Y_{ij}^{(n)}$ gives the time at which the process visits state j. The probability distribution of $Y_{ij}^{(n)}$ for fixed k is

$Y_{ij}^{(n)}$	0	1
Probability	$1 - P_{ij}^{(n)}$	$P_{ij}^{(n)}$

Thus, we have that

$$E[Y_{ij}^{(k)}] = P_{ij}^{(k)}, \quad i, j = 0, 1; \quad k = 1, 2, \ldots, n. \tag{17.3.23}$$

Because $Y_{ij}^{(k)}$ equals 1 whenever the process is in state j and 0 when it is not in j, the number of visits to j, starting originally from i, in n steps is

$$N_{ij}^{(n)} = Y_{ij}^{(1)} + Y_{ij}^{(2)} + \cdots + Y_{ij}^{(n)}. \tag{17.3.24}$$

Taking the expected values and using the property that the expectation of a sum is the sum of expectations,

$$\mu_{ij}^{(n)} = E\left[N_{ij}^{(n)}\right] = P_{ij}^{(1)} + P_{ij}^{(2)} + \cdots + P_{ij}^{(n)} = \sum_{k=1}^{n} P_{ij}^{(k)}. \qquad (\mathbf{17.3.25})$$

From Equation 17.3.13, we substitute for each $P_{ij}^{(k)}$ and find

$$\mu_{00}^{(n)} = \sum_{k=1}^{n} P_{00}^{(k)} = \sum_{k=1}^{n} \left[\frac{b}{a+b} + \frac{a(1-a-b)^k}{a+b}\right], \qquad (\mathbf{17.3.26})$$

$$\mu_{01}^{(n)} = \sum_{k=1}^{n} P_{01}^{(k)} = \sum_{k=1}^{n} \left[\frac{a}{a+b} - \frac{a(1-a-b)^k}{a+b}\right], \qquad (\mathbf{17.3.27})$$

$$\mu_{10}^{(n)} = \sum_{k=1}^{n} P_{10}^{(k)} = \sum_{k=1}^{n} \left[\frac{b}{a+b} - \frac{b(1-a-b)^k}{a+b}\right], \qquad (\mathbf{17.3.28})$$

and

$$\mu_{11}^{(n)} = \sum_{k=1}^{n} P_{11}^{(k)} = \sum_{k=1}^{n} \left[\frac{a}{a+b} + \frac{b(1-a-b)^k}{a+b}\right], \qquad (\mathbf{17.3.29})$$

finally, noting that

$$\sum_{k=1}^{n} \frac{b}{a+b} = \frac{nb}{a+b}, \qquad (\mathbf{17.3.30})$$

and

$$\sum_{k=1}^{n} \frac{a(1-a-b)^k}{a+b} = \frac{a}{a+b} \sum_{k=1}^{n} (1-a-b)^k \qquad (\mathbf{17.3.31})$$

$$= \frac{a}{a+b}\left[(1-a-b) + (1-a-b)^2 + \cdots + (1-a-b)^n\right] \qquad (\mathbf{17.3.32})$$

$$= \frac{a(1-a-b)}{a+b}\left[1 + (1-a-b) + \cdots + (1-a-b)^{n-1}\right] \qquad (\mathbf{17.3.33})$$

$$= \frac{a(1-a-b)\left[1-(1-a-b)^n\right]}{(a+b)\left[1-(1-a-b)\right]}. \qquad (\mathbf{17.3.34})$$

Here we used the property of a geometric series that

$$\sum_{k=0}^{n-1} x^k = \frac{1-x^n}{1-x}, \qquad |x| < 1. \qquad (\mathbf{17.3.35})$$

• Example 17.3.2

Let us continue with our precipitation model that we introduced in Example 17.3.1. If we wish to know the expected number of days within a week that the atmosphere will be in a given state, we have from Equation 17.3.21 that

$$\mu_{00}^{(7)} = \frac{7b}{a+b} + \frac{a(1-a-b)[1-(1-a-b)^7]}{(a+b)^2} = 3.3953, \qquad (\mathbf{17.3.36})$$

$$\mu_{10}^{(7)} = \frac{7b}{a+b} - \frac{b(1-a-b)[1-(1-a-b)^7]}{(a+b)^2} = 2.4031, \quad \textbf{(17.3.37)}$$

$$\mu_{01}^{(7)} = \frac{7a}{a+b} - \frac{a(1-a-b)[1-(1-a-b)^7]}{(a+b)^2} = 3.6047, \quad \textbf{(17.3.38)}$$

and

$$\mu_{11}^{(7)} = \frac{7a}{a+b} + \frac{b(1-a-b)[1-(1-a-b)^7]}{(a+b)^2} = 4.5969, \quad \textbf{(17.3.39)}$$

since $a = 0.3$ and $b = 0.2$. □

Duration of stay

In addition to computing the number of visits to a certain state, it would also be useful to know the fraction of the discrete time that a process stays in state j out of n when the process started in state i. These fractions are:

$$\lim_{n\to\infty} \frac{\mu_{00}^{(n)}}{n} = \lim_{n\to\infty} \frac{\mu_{10}^{(n)}}{n} = \pi_0, \quad \textbf{(17.3.40)}$$

and

$$\lim_{n\to\infty} \frac{\mu_{01}^{(n)}}{n} = \lim_{n\to\infty} \frac{\mu_{11}^{(n)}}{n} = \pi_1. \quad \textbf{(17.3.41)}$$

Thus, the limiting probabilities also give the fraction of time that the process spends in the two states in the long run.

If the process is in state i $(i = 0, 1)$ at some time, let us compute the number of additional time periods it stays in state i until it moves out of that state. We now want to show that this probability distribution α_i, $i = 0, 1$, is

$$P(\alpha_0 = n) = a(1-a)^n, \quad \textbf{(17.3.42)}$$

and

$$P(\alpha_1 = n) = b(1-b)^n, \quad \textbf{(17.3.43)}$$

where $n = 1, 2, 3, \ldots$. Furthermore,

$$E(\alpha_0) = (1-a)/a, \qquad E(\alpha_1) = (1-b)/b, \quad \textbf{(17.3.44)}$$

and

$$\mathrm{Var}(\alpha_0) = (1-a)/a^2, \qquad \mathrm{Var}(\alpha_1) = (1-b)/b^2, \quad \textbf{(17.3.45)}$$

where the transition probability matrix P of the Markov chain $\{X_n\}$ equals

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \quad \textbf{(17.3.46)}$$

with $|1-a-b| < 1$. Clearly a or b cannot equal to zero.

To prove this we note that at every step the process has two choices: either to stay in the same state or to move to the other state. Suppose the process is in state 0 at some time. The probability of a sequence of outcomes of the type $\{0\underbrace{0\cdots 0}_{n}1\}$ is required. Because of the property of Markov-dependence, we therefore have the realization of a Bernoulli process with n consecutive outcomes of one type followed by an outcome of the other type. Therefore, the probability distribution of α_0 is geometric with $(1-a)$ as the probability of "failure," and the distribution of α_1 is geometric with $(1-b)$ as the probability of failure. Thus, from Equation 16.6.5, we have that

$$P(\alpha_0 = n) = a(1-a)^n, \tag{17.3.47}$$

and

$$P(\alpha_1 = n) = b(1-b)^n, \tag{17.3.48}$$

where $n = 0, 1, 2, \ldots$. The expression for the mathematical expectation and variance of α_0 and α_1 easily follow from the corresponding expressions for the geometric distribution.

• Example 17.3.3

Let us illustrate our expectation and variance expressions for our precipitation model. From Equation 17.3.44 and Equation 17.3.45, we have that

$$E(\alpha_0) = (1-a)/a = 2.3333, \qquad E(\alpha_1) = (1-b)/b = 4, \tag{17.3.49}$$

and

$$\text{Var}(\alpha_0) = (1-a)/a^2 = 7.7777, \qquad \text{Var}(\alpha_1) = (1-b)/b^2 = 20, \tag{17.3.50}$$

since $a = 0.3$ and $b = 0.2$.

Problems

1. Given

$$P = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix},$$

(a) compute P^n and (b) find $\lim_{n\to\infty} P^n$.

2. Suppose you want to model how your dog learns a new trick. Let Fido be in state 0 if he learns the new trick and in state 1 if he fails to learn the trick. Suppose that if he learns the trick, he will retain the trick. If he has yet to learn the trick, there is a probability α of him learning it with each training session. (a) Write down the transition matrix. (b) Compute $P^{(n)}$ where n is the number of training sessions. (c) What is the steady-state solution? Interpret your result. (d) Compute the expected amount of time that Fido will spend in each state during n training sessions.

17.4 BIRTH AND DEATH PROCESSES

In the previous section we considered two-state Markov chains that undergo n steps. As the time interval between each step tends to zero, the Markov process becomes continuous in time. In this section and the next, we consider two independent examples of continuous Markov processes.

We began Chapter 16 by showing that the deterministic description of birth and death is inadequate to explain the extinction of species. Here we will fill out the details of our analysis and extend them to population dynamics and chemical kinetics. As Section 1.2

showed, deterministic models lead to first-order ordinary differential equations, and this description fails when the system initially contains a small number of particles.

Consider a population of organisms that multiply by the following rules:

1. The sub-populations generated by two co-existing individuals develop completely independently of one another;
2. an individual existing at time t has a chance $\lambda\, dt + o(dt)$ of multiplying by binary fission during the following time interval of length dt;
3. the "birth rate" λ is the same for all individuals in the population at any time t;
4. an individual existing at time t has a chance $\mu\, dt + o(dt)$ of dying in the following time interval of length dt; and
5. the "death rate" μ is the same for all individuals at any time t.

Rule 3 is usually interpreted in the sense that in each birth just one new member is added to the population, but of course mathematically (and because the age structure of the population is being ignored) it is not possible to distinguish between this and an alternative interpretation in which one of the parents dies when the birth occurs and is replaced by two children.

Let n_0 be the number of individuals at the initial time $t = 0$ and let $p_n(t)$ denote the probability that the population size $N(t)$ has the value n at the time t. Then

$$\frac{dp_n}{dt} = (n-1)\lambda p_{n-1} - n(\lambda+\mu)p_n + \mu(n+1)p_{n+1}, \qquad n \geq 1, \tag{17.4.1}$$

and

$$\frac{dp_0(t)}{dt} = \mu p_1(t), \tag{17.4.2}$$

subject to the initial condition that

$$p_n(0) = \begin{cases} 1, & n = n_0, \\ 0, & n \neq n_0. \end{cases} \tag{17.4.3}$$

Equation 17.4.1 through Equation 17.4.3 constitute a system of linear ordinary equations. The question now turns on how to solve them most efficiently. To this end we introduce a probability-generating function:

$$\phi(z,t) = \sum_{n=0}^{\infty} z^n p_n(t). \tag{17.4.4}$$

Summing Equation 17.4.1 from $n = 1$ to ∞ after we multiplied it by z^n and using Equation 17.4.2, we obtain

$$\sum_{n=0}^{\infty} z^n \frac{dp_n}{dt} = \lambda \sum_{n=1}^{\infty}(n-1)z^n p_{n-1}(t) - (\lambda+\mu)\sum_{n=1}^{\infty} nz^n p_n(t) + \mu \sum_{n=0}^{\infty}(n+1)z^n p_{n+1}(t). \tag{17.4.5}$$

Because

$$\sum_{n=0}^{\infty} z^n \frac{dp_n}{dt} = \frac{\partial \phi}{\partial t}, \tag{17.4.6}$$

$$\sum_{n=1}^{\infty} nz^n p_n(t) = z\sum_{n=0}^{\infty} nz^{n-1} p_n(t) = z\frac{\partial \phi}{\partial z}, \tag{17.4.7}$$

$$\sum_{n=1}^{\infty}(n-1)z^n p_{n-1}(t) = \sum_{k=0}^{\infty} k z^{k+1} p_k(t) = z^2 \sum_{k=0}^{\infty} k z^{k-1} p_k(t) = z^2 \frac{\partial \phi}{\partial z}, \tag{17.4.8}$$

and

$$\sum_{n=0}^{\infty}(n+1)z^n p_{n+1}(t) = \sum_{k=1}^{\infty} k z^{k-1} p_k(t) = \sum_{k=0}^{\infty} k z^{k-1} p_k(t) = \frac{\partial \phi}{\partial z}, \tag{17.4.9}$$

Equation 17.4.5 becomes the first-order partial differential equation

$$\frac{\partial \phi}{\partial t} = (\lambda z - \mu)(z-1)\frac{\partial \phi}{\partial z}, \tag{17.4.10}$$

subject to the initial condition

$$\phi(z,0) = z^{n_0}. \tag{17.4.11}$$

Equation 17.4.10 is an example of a first-order partial differential equation of the general form

$$P(x,y)\frac{\partial u}{\partial x} + Q(x,y)\frac{\partial u}{\partial y} = 0. \tag{17.4.12}$$

This equation has solutions[4] of the form $u(x,y) = f(\xi)$ where $f(\cdot)$ is an arbitrary function that is differentiable and $\xi(x,y) =$ constant are solutions to

$$\frac{dx}{P(x,y)} = \frac{dy}{Q(x,y)}. \tag{17.4.13}$$

In the present case,

$$\frac{dt}{1} = -\frac{dz}{(\lambda z - \mu)(z-1)} = -\frac{dz}{(\lambda - \mu)(z-1)} + \frac{dz}{(\lambda - \mu)(z - \mu/\lambda)}. \tag{17.4.14}$$

Integrating Equation 17.4.14,

$$-(\lambda - \mu)t + \ln[\psi(z)] = \ln(\xi), \tag{17.4.15}$$

or

$$\xi(z,t) = \psi(z)e^{-(\lambda-\mu)t}, \tag{17.4.16}$$

where

$$\psi(z) = \frac{\lambda z - \mu}{z - 1}. \tag{17.4.17}$$

Therefore, the general solution is

$$\phi(z,t) = f\left[\psi(z)e^{-(\lambda-\mu)t}\right]. \tag{17.4.18}$$

Our remaining task is to find $f(\cdot)$. From the initial condition, Equation 17.4.11, we have that

$$\phi(z,0) = f[\psi(z)] = z^{n_0}. \tag{17.4.19}$$

---

[4] See Webster, A. G., 1966: *Partial Differential Equations of Mathematical Physics*. Dover, 446 pp. See Section 22.

Because $z = [\mu - \psi(z)]/[\lambda - \psi(z)]$, then

$$f(\psi) = \left(\frac{\mu - \psi}{\lambda - \psi}\right)^{n_0}. \qquad \textbf{(17.4.20)}$$

Therefore,

$$\phi(z,t) = \left[\frac{\mu - \psi(z)e^{-(\lambda-\mu)t}}{\lambda - \psi(z)e^{-(\lambda-\mu)t}}\right]^{n_0}. \qquad \textbf{(17.4.21)}$$

Once we find $\phi(z,t)$, we can compute the probabilities of each of the species from the probability generating function. For example,

$$P\{N(t) = 0|N(0) = n_0\} = p_0(t) = \phi(0,t). \qquad \textbf{(17.4.22)}$$

From Equation 17.4.17 we have $\psi(0) = \mu$ and

$$\phi(0,t) = \left\{\frac{\mu\left[1 - e^{-(\lambda-\mu)t}\right]}{\lambda - \mu e^{-(\lambda-\mu)t}}\right\}^{n_0}, \qquad \lambda \neq \mu, \qquad \textbf{(17.4.23)}$$

and

$$\phi(0,t) = \left(\frac{\lambda t}{1 + \lambda t}\right)^{n_0}, \qquad \lambda = \mu. \qquad \textbf{(17.4.24)}$$

An important observation from Equation 17.4.23 and Equation 17.4.24 is that

$$\lim_{t\to\infty} p_0(t) = 1, \qquad \lambda \leq \mu, \qquad \textbf{(17.4.25)}$$

and

$$\lim_{t\to\infty} p_0(t) = \left(\frac{\mu}{\lambda}\right)^{n_0}, \qquad \lambda > \mu. \qquad \textbf{(17.4.26)}$$

This limit can be interpreted as the probability of extinction of the population in a finite time. Consequently there will be "almost certain" extinction whenever $\lambda \leq \mu$. These results, which are true whatever the initial number of individuals may be, show very clearly the inadequacy of the deterministic description of population dynamics.

Finally, let us compute the mean and variance for the birth and death process. The expected number of individuals at time t is

$$m(t) = E[N(t)] = \sum_{n=0}^{\infty} n\, p_n(t) = \sum_{n=1}^{\infty} n\, p_n(t). \qquad \textbf{(17.4.27)}$$

Now

$$\frac{dm}{dt} = \sum_{n=1}^{\infty} n \frac{dp_n}{dt} = \sum_{n=1}^{\infty} n\left[(n-1)\lambda p_{n-1} - n(\lambda+\mu)p_n + \mu(n+1)p_{n+1}\right] \qquad \textbf{(17.4.28)}$$

$$= \lambda \sum_{n=1}^{\infty}(n-1)^2 p_{n-1} + \lambda \sum_{n=1}^{\infty}(n-1)p_{n-1} - (\lambda+\mu)\sum_{n=1}^{\infty} n^2 p_n + \mu \sum_{n=1}^{\infty}(n+1)^2 p_{n+1}$$

$$- \mu \sum_{n=1}^{\infty}(n+1)p_{n+1} \qquad \textbf{(17.4.29)}$$

$$= -(\lambda+\mu)\sum_{n=1}^{\infty} n^2 p_n + \lambda \sum_{i=0}^{\infty} i^2 p_m + \mu \sum_{k=2}^{\infty} k^2 p_k + \lambda \sum_{i=0}^{\infty} i\, p_i - \mu \sum_{k=2}^{\infty} k\, p_k. \qquad \textbf{(17.4.30)}$$

In the first three sums in Equation 17.4.30, terms from $i, k, n = 2$, and onward cancel and leave $-(\lambda+\mu)p_1 + \lambda p_1 = -\mu p_1$. Therefore,

$$\frac{dm}{dt} = -\mu p_1 + \lambda \sum_{i=0}^{\infty} i\, p_i - \mu \sum_{k=2}^{\infty} k\, p_k = (\lambda - \mu) \sum_{n=0}^{\infty} n\, p_n, = (\lambda-\mu)m. \tag{17.4.31}$$

If we choose the initial condition $m(0) = n_0$, the solution is

$$m(t) = n_0 e^{(\lambda-\mu)t}. \tag{17.4.32}$$

This is the same as the deterministic result, Equation 1.2.9, with the birth rate $\overline{b}$ replaced by λ and the death rate $\overline{d}$ replaced by μ. Furthermore, if $\lambda = \mu$, the mean size of the population is constant.

The second moment of $N(t)$ is

$$M(t) = \sum_{n=0}^{\infty} n^2 p_n(t). \tag{17.4.33}$$

Proceeding as before, we have that

$$\frac{dM}{dt} = \sum_{n=1}^{\infty} n^2 \frac{dp_n}{dt} = \sum_{n=1}^{\infty} n^2 \left[\lambda(n-1)p_{n-1} - (\lambda+\mu)n\, p_n + \mu(n+1)p_{n+1}\right] \tag{17.4.34}$$

$$= \lambda \sum_{n=1}^{\infty}(n-1)^3 p_{n-1} + 2\lambda \sum_{n=1}^{\infty}(n-1)^2 p_{n-1} + \lambda \sum_{n=1}^{\infty}(n-1)p_{n-1} - (\lambda+\mu)\sum_{n=1}^{\infty} n^3 p_n$$
$$+ \mu \sum_{n=1}^{\infty}(n+1)^3 p_{n+1} - 2\mu \sum_{n=1}^{\infty}(n+1)^2 p_{n+1} + \mu \sum_{n=1}^{\infty}(n+1)p_{n+1} \tag{17.4.35}$$

$$= \lambda \sum_{k=1}^{\infty} k^3 p_k + 2\lambda \sum_{k=1}^{\infty} k^2 p_k + \lambda \sum_{k=1}^{\infty} k\, p_k - (\lambda+\mu)\sum_{n=1}^{\infty} n^3 p_n$$
$$+ \mu \sum_{i=2}^{\infty} i^3 p_i - 2\mu \sum_{i=2}^{\infty} i^2 p_i + \mu \sum_{i=2}^{\infty} i\, p_i. \tag{17.4.36}$$

The three sums, which contain either i^3 or k^3 or n^3 in them, cancel when $i, k, n = 2$ and onward; these three sums reduce to $-\mu p_1$. The sums that involve i^2 or k^2 can be written in terms of $M(t)$. Finally, the sums involving i and k can be expressed in terms of $m(t)$. Therefore, Equation 17.4.36 becomes the first-order ordinary differential equation

$$\frac{dM}{dt} - 2(\lambda-\mu)M = (\lambda+\mu)m(t) = (\lambda+\mu)n_0 e^{(\lambda-\mu)t}, \tag{17.4.37}$$

with $M(0) = n_0^2$.

Equation 17.4.37 can be solved exactly using the technique of integrating factors from Section 1.5. Its solution is

$$M(t) = n_0^2 e^{2(\lambda-\mu)t} + \frac{\lambda+\mu}{\lambda-\mu} n_0 e^{(\lambda-\mu)t}\left[e^{(\lambda-\mu)t} - 1\right]. \tag{17.4.38}$$

From the definition of variance, Equation 16.6.5, the variance of the population in the birth and death process equals

$$\text{Var}[N(t)] = n_0 \frac{(\lambda+\mu)}{(\lambda-\mu)} e^{(\lambda-\mu)t}\left[e^{(\lambda-\mu)t} - 1\right], \qquad \lambda \neq \mu, \tag{17.4.39}$$

or

$$\text{Var}[N(t)] = 2\lambda n_0 t, \qquad \lambda = \mu. \tag{17.4.40}$$

• Example 17.4.1: Chemical kinetics

The use of Markov processes to describe birth and death has become quite popular. Indeed, it can be applied to any phenomena where something is being created or destroyed. Here we illustrate its application in chemical kinetics.

Let the random variable $X(t)$ be the number of A molecules in a unimolecular reaction $A \to B$ (such as radioactive decay) at time t. A stochastic model that describes the decrease of A can be constructed from the following assumptions:

1. The probability of transition from n to $n-1$ in the time interval $(t, t+\Delta t)$ is $n\lambda\Delta t + o(\Delta t)$ where λ is a constant and $o(\Delta t)$ denotes that $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$.
2. The probability of a transition from n to $n-j$, $j > 1$, in the time interval $(t, t+\Delta t)$ is at least $o(\Delta t)$ because the time interval is so small that only one molecule undergoes a transition.
3. The reverse reaction occurs with probability zero.

The equation that governs the probability that $X(t) = n$ is

$$p_n(t + \Delta t) = (n+1)\lambda\Delta t p_{n+1}(t) + (1 - \lambda n\Delta t)p_n(t) + o(\Delta t). \tag{17.4.41}$$

Transposing $p_n(t)$ from the right side, dividing by Δt, and taking the limit $\Delta t \to 0$, we obtain the differential-difference equation[5]

$$\frac{dp_n}{dt} = (n+1)\lambda p_{n+1}(t) - n\lambda p_n(t). \tag{17.4.42}$$

Equation 17.4.42 is frequently called the *stochastic master equation*. The first term on the right side of this equation vanishes when $n = n_0$.

The solution of Equation 17.4.42 once again involves introducing a generating function for $p_n(t)$, namely

$$F(z,t) = \sum_{n=0}^{n_0} p_n(t)z^n, \qquad |z| < 1. \tag{17.4.43}$$

Summing Equation 17.4.42 from $n = 0$ to n_0 after multiplying it by z^n, we find

$$\sum_{n=0}^{n_0} z^n \frac{dp_n}{dt} = \lambda \sum_{n=0}^{n_0-1} (n+1)z^n p_{n+1}(t) - \lambda \sum_{n=1}^{n_0} nz^n p_n(t). \tag{17.4.44}$$

Because

$$\sum_{n=0}^{n_0} z^n \frac{dp_n}{dt} = \frac{\partial F}{\partial t}, \tag{17.4.45}$$

$$\sum_{n=0}^{n_0} nz^n p_n(t) = z\sum_{n=0}^{n_0} nz^{n-1} p_n(t) = z\frac{\partial F}{\partial z}, \tag{17.4.46}$$

---

[5] McQuarrie, D. A., 1963: Kinetics of small systems. I. *J. Chem. Phys.*, **38**, 433–436.

and

$$\sum_{n=0}^{n_0-1}(n+1)z^n p_{n+1}(t) = \sum_{k=1}^{n_0} kz^{k-1}p_k(t) = \sum_{k=0}^{n_0} kz^{k-1}p_k(t) = \frac{\partial F}{\partial z}, \tag{17.4.47}$$

Equation 17.4.44 becomes the first-order partial differential equation

$$\frac{\partial F}{\partial t} = \lambda(1-z)\frac{\partial F}{\partial z}. \tag{17.4.48}$$

The solution of Equation 17.4.48 follows the method used to solve Equation 17.4.10. Here we find $\xi(z,t)$ via

$$\frac{dt}{1} = \frac{dz}{\lambda(z-1)}, \tag{17.4.49}$$

or

$$\xi(z,t) = (z-1)e^{-\lambda t}. \tag{17.4.50}$$

Therefore,

$$F(z,t) = f\left[(z-1)e^{-\lambda t}\right]. \tag{17.4.51}$$

To find $f(\cdot)$, we use the initial condition that $F(z,0) = z^{n_0}$. This yields $f(y) = (1+y)^{n_0}$ and

$$F(z,t) = \left[1+(z-1)e^{-\lambda t}\right]^{n_0}. \tag{17.4.52}$$

Once again, we can compute the mean and variance of this process. Because

$$\left.\frac{\partial F}{\partial z}\right|_{z=1} = \sum_{n=0}^{n_0} np_n(t), \tag{17.4.53}$$

the mean is given by

$$E[X(t)] = \frac{\partial F(1,t)}{\partial z}. \tag{17.4.54}$$

To compute the variance, we first compute the second moment. Since

$$z\frac{\partial F}{\partial z} = \sum_{n=0}^{n_0} nz^n p_n(t), \tag{17.4.55}$$

and

$$\frac{\partial}{\partial z}\left(z\frac{\partial F}{\partial z}\right) = \sum_{n=0}^{n_0} n^2 z^{n-1} p_n(t), \tag{17.4.56}$$

we have that

$$\sum_{n=0}^{n_0} n^2 p_n(t) = \frac{\partial^2 F(1,t)}{\partial z^2} + \frac{\partial F(1,t)}{\partial z}. \tag{17.4.57}$$

From Equation 16.6.5, the final result is

$$\text{Var}[X(t)] = \frac{\partial^2 F(1,t)}{\partial z^2} + \frac{\partial F(1,t)}{\partial z} - \left[\frac{\partial F(1,t)}{\partial z}\right]^2. \tag{17.4.58}$$

Upon substituting Equation 17.4.52 into Equations 17.4.54 and 17.4.58, the mean and variance for this process are

$$E[X(t)] = n_0 e^{-\lambda t}, \qquad \text{and} \qquad \text{Var}[X(t)] = n_0 e^{-\lambda t}\left(1-e^{-\lambda t}\right). \tag{17.4.59}$$

Because the expected value of the stochastic representation also equals the deterministic result, the two representations are "consistent in the mean." Further study shows that this is true only for unimolecular reactions. Upon expanding Equation 17.4.52, we find that

$$p_n(t) = \binom{n_0}{n} e^{-n\lambda t} \left(1 - e^{-\lambda t}\right)^{n_0 - n}. \tag{17.4.60}$$

An alternative method to the generating function involves Laplace transforms.[6] To illustrate this method, we again examine the reaction $A \to B$. The stochastic master equation is

$$\frac{dp_n}{dt} = (n-1)\lambda p_{n-1}(t) - n\lambda p_n(t), \qquad n_0 \le n < \infty, \tag{17.4.61}$$

$p_n(t) = 0$ for $0 < n < n_0$, where $p_n(t)$ denotes the probability that we have n particles of B at time t. The initial condition is that $p_{n_0}(0) = 1$ and $p_m(0) = 0$ for $m \neq n_0$ where n_0 denotes the initial number of molecules of B.

Taking the Laplace transform of Equation 17.4.61, we find that

$$sP_n(s) = (n-1)\lambda P_{n-1}(s) - n\lambda P_n(s), \qquad n_0 < n < \infty, \tag{17.4.62}$$

and

$$sP_{n_0}(s) - 1 = -n\lambda P_{n_0}(s). \tag{17.4.63}$$

Therefore, solving for $P_n(s)$,

$$P_n(s) = \frac{(n-1)\lambda}{s + n\lambda} P_{n-1}(s) = \frac{\lambda^{n-n_0}(n-1)!}{(n_0-1)!} \prod_{k=n_0}^{n} (s + k\lambda)^{-1}. \tag{17.4.64}$$

From partial fractions,

$$P_n(s) = \frac{(n-1)!}{(n_0-1)!} \sum_{k=n_0}^{n} \frac{(-1)^{k-n_0}}{(k-n_0)!(n-k)!(s+k\lambda)}. \tag{17.4.65}$$

Taking the inverse Laplace transform,

$$p_n(t) = \frac{(n-1)!}{(n_0-1)!(n-n_0)!} \sum_{k=n_0}^{n} \frac{(-1)^{k-n_0}(n-n_0)!}{(k-n_0)!(n-k)!} e^{-\lambda kt} \tag{17.4.66}$$

$$= \frac{(n-1)!e^{-\lambda n_0 t}}{(n_0-1)!(n-n_0)!} \sum_{j=0}^{n-n_0} \frac{(-1)^j (n-n_0)!}{j!(n-n_0-j)!} e^{-\lambda jt} \tag{17.4.67}$$

$$= \frac{(n-1)!e^{-\lambda n_0 t}}{(n_0-1)!(n-n_0)!} \left(1 - e^{-\lambda t}\right)^{n-n_0}, \tag{17.4.68}$$

where we introduced $j = k - n_0$ and eliminated the summation via the binomial theorem. Equation 17.4.68 is identical to results[7] given by Delbrück using another technique. □

---

[6] Ishida, K., 1969: Stochastic model for autocatalytic reaction. *Bull. Chem. Soc. Japan*, **42**, 564–565.

[7] Delbrück, M., 1940: Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.*, **8**, 120–124. See his Equation 7.

• Example 17.4.2

In the chemical reaction $rA \overset{\lambda}{\underset{\mu}{\rightleftarrows}} B$, r molecules of A combine to form one molecule of B. If $X(t) = n$ is the number of B molecules, then the probability $p_n(t) = P\{X(t) = n\}$ of having n molecules of B is given by

$$\frac{dp_n}{dt} = -\left[n\mu + (N-n)\lambda\right] p_n + (N-n+1)\lambda p_{n-1} + (n+1)\mu p_{n+1}, \tag{17.4.69}$$

where $0 \le n \le N$, rN is the total number of molecules of A, λ is the rate at which r molecules of A combine to produce B, and μ is the rate at which B decomposes into A.

Multiplying Equation 17.4.69 by z^n and summing from $n = -1$ to $N+1$,

$$\begin{aligned}\sum_{n=-1}^{N+1} z^n \frac{dp_n}{dt} &= -N\lambda \sum_{n=-1}^{N+1} z^n p_n + (\lambda-\mu) \sum_{n=-1}^{N+1} nz^n p_n + N\lambda \sum_{n=-1}^{N+1} z^n p_{n-1} \\ &\quad - \lambda \sum_{n=-1}^{N+1} (n-1) z^n p_{n-1} + \mu \sum_{n=-1}^{N+1} (n+1) z^n p_{n+1}.\end{aligned} \tag{17.4.70}$$

Defining

$$F(z,t) = \sum_{n=-1}^{N+1} p_n(t) z^n, \qquad |z| < 1, \tag{17.4.71}$$

with $p_{-1} = p_{N+1} = 0$, we have that

$$\frac{\partial F}{\partial t} = \sum_{n=-1}^{N+1} z^n \frac{dp_n}{dt}, \tag{17.4.72}$$

$$\frac{\partial F}{\partial z} = \sum_{n=-1}^{N+1} nz^{n-1} p_n = \sum_{i=-2}^{N} (i+1) z^i p_{i+1} = \sum_{i=-1}^{N+1} (i+1) z^i p_{i+1}, \tag{17.4.73}$$

$$z\frac{\partial F}{\partial z} = \sum_{n=-1}^{N+1} nz^n p_n, \tag{17.4.74}$$

$$z^2 \frac{\partial F}{\partial z} = \sum_{n=-1}^{N+1} nz^{n+1} p_n = \sum_{i=0}^{N+2} (i-1) z^i p_{i-1} = \sum_{i=-1}^{N+1} (i-1) z^i p_{i-1}, \tag{17.4.75}$$

and

$$F = \sum_{n=-1}^{N+1} z^{n+1} p_n = \sum_{i=0}^{N+2} z^i p_{i-1} = \sum_{i=-1}^{N+1} z^i p_{i-1}. \tag{17.4.76}$$

Therefore, the differential-difference equation, Equation 17.4.70, can be replaced by

$$\frac{\partial F}{\partial t} = N\lambda(z-1)F + \left[\mu - (\mu-\lambda)z - \lambda z^2\right] \frac{\partial F}{\partial z}. \tag{17.4.77}$$

Using the same technique as above, this partial differential equation can be written as

$$\frac{dt}{-1} = \frac{dz}{(1-z)(\mu+\lambda z)} = \frac{dF}{-N\lambda(z-1)}. \tag{17.4.78}$$

Equation 17.4.78 yields the independent solutions

$$\frac{1-z}{\mu+\lambda z} e^{-(\mu+\lambda)t} = \xi(z,t) = \text{constant}, \qquad \textbf{(17.4.79)}$$

and

$$(\mu+\lambda)^{-N} F(z,t) = \eta(z,t) = \text{another constant}, \qquad \textbf{(17.4.80)}$$

where $f(\cdot)$ is an arbitrary, differentiable function. If there are m units of B at $t = 0$, $0 \le m \le N$, the initial condition is $F(z,0) = z^m$. Then,

$$f\left(\frac{1-z}{\mu+\lambda z}\right) = \frac{z^m}{(\mu+\lambda z)^N}, \qquad \textbf{(17.4.81)}$$

or

$$f(x) = \frac{(1-\mu x)^m}{(\mu+\lambda)^N} (1+\lambda x)^{N-m}. \qquad \textbf{(17.4.82)}$$

After some algebra, we finally find that

$$F(z,t) = \frac{1}{(\mu+\lambda)^N} \left\{\mu\left[1-e^{-(\mu+\lambda)t}\right] + z\left[\lambda + \mu e^{-(\mu+\lambda)t}\right]\right\}^m \times \left\{\mu + \lambda e^{-(\mu+\lambda)t} + \lambda z\left[1-e^{-(\mu+\lambda)t}\right]\right\}^{N-m}. \qquad \textbf{(17.4.83)}$$

Computing the mean and variance, we obtain

$$E(X) = \frac{m}{\mu+\lambda}\left[\lambda + \mu e^{-(\mu+\lambda)t}\right] + \frac{(N-m)\lambda}{\mu+\lambda}\left[1-e^{-(\mu+\lambda)t}\right], \qquad \textbf{(17.4.84)}$$

and

$$\text{Var}(X) = \frac{m\mu}{(\mu+\lambda)^2}\left[\lambda+\mu e^{-(\mu+\lambda)t}\right]\left[1-e^{-(\mu+\lambda)t}\right] + \frac{(N-m)\lambda}{(\mu+\lambda)^2}\left[\mu+\lambda e^{-(\mu+\lambda)t}\right]\left[1-e^{-(\mu+\lambda)t}\right]. \qquad \textbf{(17.4.85)}$$

Problems

1. During their study of growing cancerous cells (with growth rate α), Bartoszyński et al.[8] developed a probabilistic model of a tumor that has not yet metastasized. In their mathematical derivation a predictive model gives the probability $p_n(t)$ that certain nth type of cells (out of N) will develop. This probability can change in two ways: (1) Each of the existing cells has the probability $\lambda n \Delta t + o(\Delta t)$ of mutating to another type between t and $t+\Delta t$. (2) The probability that cells in state n at time t will shed a metastasis between t and $t+\Delta t$ is $\mu n c e^{t/\alpha}\Delta t + o(\Delta t)$, where μ is a constant and c is the size of a single cell. Setting $\rho = \lambda c/N$ and $\nu = \mu c$, the governing equations for $p_n(t)$ is

$$\frac{dp_n}{dt} = -(\rho+\nu) n e^{t/\alpha} p_n + \rho(n+1) e^{t/\alpha} p_{n+1}, \qquad n = 0,1,2,\ldots,N-1, \qquad (1)$$

[8] Bartoszyński, R., B. F. Jones, and J. P. Klein, 1985: Some stochastic models of cancer metastases. *Commun. Statist.-Stochastic Models*, **1**, 317–339.

and

$$\frac{dp_N}{dt} = -(\rho + \nu)Ne^{t/\alpha}p_N, \tag{2}$$

with the initial conditions $p_N(0) = 1$ and $p_n(0) = 0$ if $n \neq N$.

Step 1: Introducing the generating function

$$\phi(z,t) = \sum_{n=0}^{N} z^n p_n(t), \qquad 0 \leq z \leq 1, \tag{3}$$

show that our system of linear differential-difference equations can be written as the first-order partial differential equation

$$\frac{\partial \phi}{\partial t} = [\rho - (\rho + \nu)z]e^{t/\alpha}\frac{\partial \phi}{\partial z} \tag{4}$$

with $\phi(z,0) = z^N$.

Step 2: Solve the partial differential equation in Step 1 and show that

$$\phi(z,t) = \left(\frac{\rho}{\rho+\nu}\right)^N \left\{1 - \left(1 - \frac{\rho+\nu}{\rho}z\right)\exp\left[-\alpha(\rho+\nu)\left(e^{t/a} - 1\right)\right]\right\}^N. \tag{5}$$

Project: Stochastic Simulation of Chemical Reactions

Most stochastic descriptions of chemical reactions cannot be attacked analytically and numerical simulation is necessary. The purpose of this project is to familiarize you with some methods used in the stochastic simulation of chemical reactions. In particular, we will use the Lokta reactions given by the reaction equations:

$$A + X \xrightarrow{k_1} 2X, \tag{1}$$

$$X + Y \xrightarrow{k_2} 2Y, \tag{2}$$

$$Y \xrightarrow{k_3} Z. \tag{3}$$

Surprisingly, simple numerical integration of the master equation is not fruitful. This occurs because of the number and nature of the independent variables; there is only one master equation but N reactants and time for independent variables.

An alternative to integrating the master equation is a direct stochastic simulation. In this approach, the (transition) probability for each reaction is computed: $p_1 = k_1 ax\Delta t$, $p_2 = k_2 xy\Delta t$, and $p_3 = k_3 y\Delta t$, where Δt is the time between each consecutive state and a is the constant number of molecules of A. The obvious question is: Which of these probabilities should we use?

Our first attempt follows Nakanishi:[9] Assume that Δt is sufficiently small so that $p_1 + p_2 + p_3 < 1$. Using a normalized uniform distribution, such as MATLAB's `rand`, compute a random variable `r` for each time step. Then march forward in time. At each time step, there are four possibilities. If $0 < \mathtt{r} \leq p_1$, then the first reaction occurs and

[9] This is the technique used by Nakanishi, T., 1972: Stochastic analysis of an oscillating chemical reaction. *J. Phys. Soc. Japan*, **32**, 1313–1322.

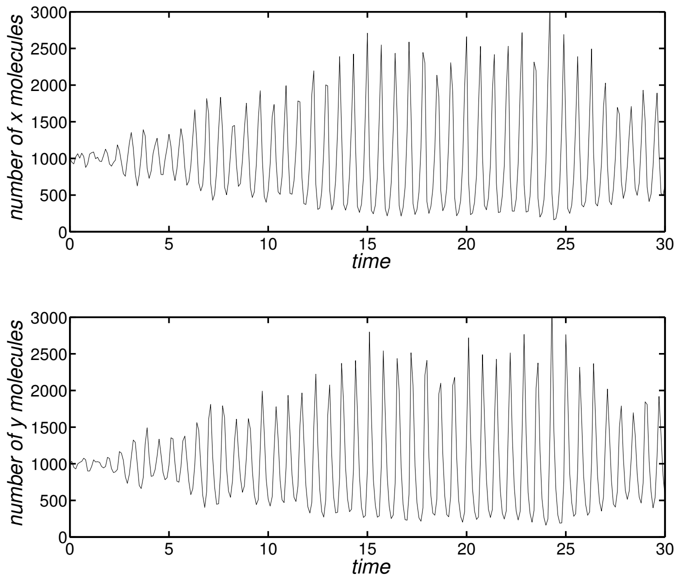


Figure 17.4.1: The temporal variation of the molecules in a Lokta reaction when $\Delta t = 10^{-5}$, $k_1 a = 10$, $k_2 = 0.01$, $k_3 = 10$, and $x(0) = y(0) = 1000$.

$x(t+\Delta t) = x(t)+1$, $y(t+\Delta t) = y(t)$. If $p_1 < \mathtt{r} \le p_1 + p_2$, then the second reaction occurs and $x(t+\Delta t) = x(t)-1$, $y(t+\Delta t) = y(t)+1$. If $p_1 + p_2 < \mathtt{r} \le p_1 + p_2 + p_3$, then the third reaction occurs and $x(t+\Delta t) = x(t)$, $y(t+\Delta t) = y(t) - 1$. Finally, if $p_1 + p_2 + p_3 < \mathtt{r} \le 1$, then no reaction occurs and $x(t+\Delta t) = x(t)$, $y(t+\Delta t) = y(t)$.

For the first portion of this project, create MATLAB code to simulate our chemical reaction using this simulation technique. Explore how your results behave as you vary $x(0)$, $y(0)$ and especially Δt. See Figure 17.4.1.

One of the difficulties in using Nakanishi's method is the introduction of Δt. What value should we choose to ensure that $p_1 + p_2 + p_3 < 1$? Several years later, Gillespie[10] developed a similar algorithm. He introduced three parameters, $a_1 = k_1 ax$, $a_2 = k_2 xy$, and $a_3 = k_3 y$, along with $a_0 = a_1 + a_2 + a_3$. These parameters a_1, a_2, and a_3 are similar to the probabilities p_1, p_2, and p_3. Similarly, he introduced a random number r_2 that is chosen from a normalized uniform distribution. Then, if $0 < r_2 a_0 \le a_1$, the first reaction occurs and $x(t+\Delta t) = x(t)+1$, $y(t+\Delta t) = y(t)$. If $a_1 < r_2 a_0 \le a_1 + a_2$, then the second reaction occurs and $x(t+\Delta t) = x(t)-1$, $y(t+\Delta t) = y(t)+1$. If $a_1 + a_2 < r_2 a_0 \le a_0$, then the third reaction occurs and $x(t+\Delta t) = x(t)$, $y(t+\Delta t) = y(t) - 1$. Because of his selection criteria for the reaction that occurs during a time step, one of the three reactions must take place. See Figure 17.4.2.

The most radical difference between the Nakanishi and Gillespie schemes involves the time step. It is no longer constant but varies with time and equals $\Delta t = \ln(1/r_1)/a_0$, where r_1 is a random variable selected from a normalized uniform distribution. The theoretical justification for this choice is given in Section III of his paper.

---

[10] Gillespie, D. T., 1976: A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, **22**, 403–434; Gillespie, D. T., 1977: Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, **81**, 2340–2361

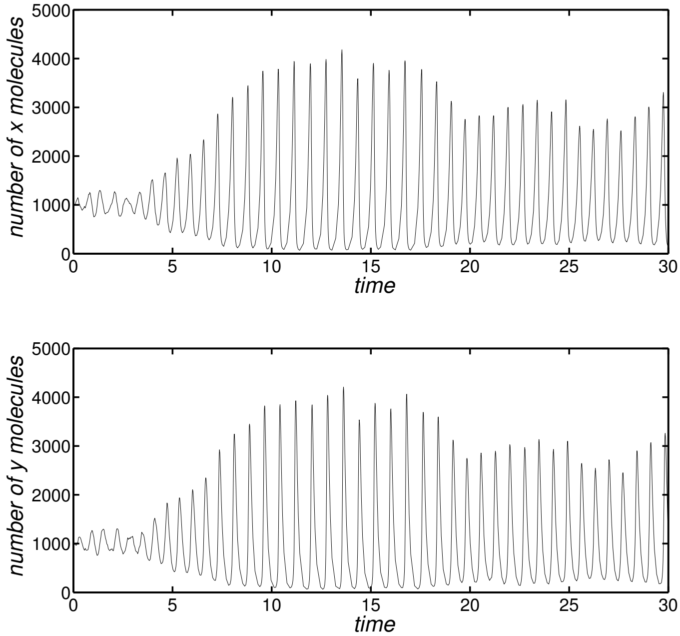


Figure 17.4.2: Same as Figure 17.4.1 except that Gillespie's method has been used.

For the second portion of this project, create MATLAB code to simulate our chemical reaction using Gillespie's technique. You might like to plot $x(t)$ vs $y(t)$ and observe the patterns that you obtain.

Finally, for a specific time, compute the probability density function that gives the probability that x and y molecules exist. See Figure 17.4.3.

17.5 POISSON PROCESSES

The Poisson random process is a counting process that counts the number of occurrences of some particular event as time increases. In other words, for each value of t, there is a number $N(t)$, which gives the number of events that occurred during the interval $[0, t]$. For this reason $N(t)$ is a discrete random variable with the set of possible values $\{0, 1, 2, \ldots\}$. Figure 17.5.1 illustrates a sample function. We can express this process mathematically by

$$N(t) = \sum_{n=0}^{\infty} H(t - T[n]), \tag{17.5.1}$$

where $T[n]$ is the time to the nth arrival, a random sequence of times. The question now becomes how to determine the values of $T[n]$. The answer involves three rather physical assumptions. They are:

1. $N(0) = 0$.
2. $N(t)$ has independent and stationary increments. By stationary we mean that for any two equal time intervals Δt_1 and Δt_2, the probability of n events in Δt_1 equals the probability of n events in Δt_2. By independent we mean that for any time interval

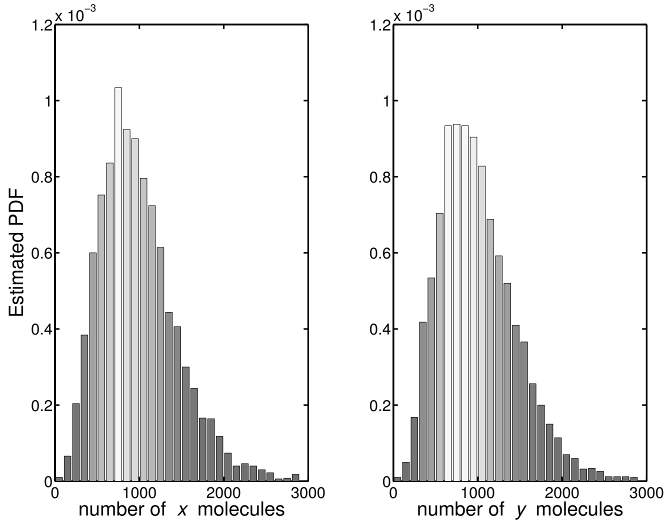


Figure 17.4.3: The estimated probability density function for the chemical reactions given by Equations (1) through (3) (for X on the left, Y on the right) at time $t = 10$. Five thousand realizations were used in these computations.

$(t, t + \Delta t)$ the probability of n events in $(t, t + \Delta t)$ is independent of how many events have occurred earlier or how they have occurred.

3.

$$P[N(t + \Delta t) - N(t) = k\,] = \begin{cases} 1 - \lambda\Delta t, & k = 0, \\ \lambda\Delta t, & k = 1, \\ 0, & k > 1, \end{cases} \qquad \mathbf{(17.5.2)}$$

for all t. Here λ equals the expected number of events in an interval of unit length of time. Because $E[N(t)] = \lambda$, it is the average number of events that occur in one unit of time and in practice it can be measured experimentally.

We begin our analysis of Poisson processes by finding $P[N(t) = 0\,]$ for any $t > 0$. If there are no arrivals in $[0, t]$, then there must be no arrivals in $[0, t - \Delta t]$ and also no arrivals in $(t - \Delta t, t]$. Therefore,

$$P[N(t) = 0\,] = P[N(t - \Delta t) = 0, N(t) - N(t - \Delta t) = 0\,]. \qquad \mathbf{(17.5.3)}$$

Because $N(t)$ is independent,

$$P[N(t) = 0\,] = P[N(t - \Delta t) = 0\,]P[N(t) - N(t - \Delta t) = 0\,]. \qquad \mathbf{(17.5.4)}$$

Furthermore, since $N(t)$ is stationary,

$$P[N(t) = 0\,] = P[N(t - \Delta t) = 0\,]P[N(t + \Delta t) - N(t) = 0\,]. \qquad \mathbf{(17.5.5)}$$

Finally, from Equation 17.5.2,

$$P[N(t) = 0\,] = P[N(t - \Delta t) = 0\,](1 - \lambda\Delta t). \qquad \mathbf{(17.5.6)}$$

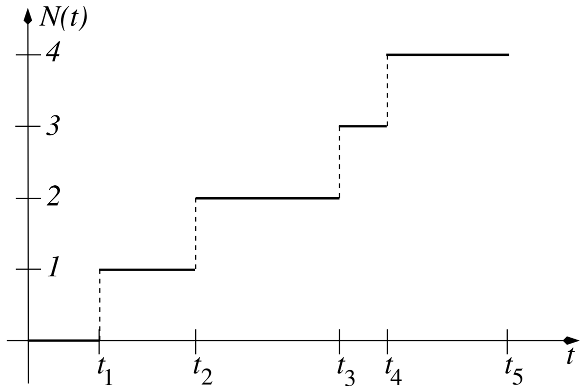


Figure 17.5.1: Schematic of a Poisson process.

Let us denote $P[N(t)=0\,]$ by $P_0(t)$. Then,

$$P_0(t) = P_0(t-\Delta t)(1-\lambda\Delta t), \tag{17.5.7}$$

or

$$\frac{P_0(t)-P_0(t-\Delta t)}{\Delta t} = -\lambda P_0(t-\Delta t). \tag{17.5.8}$$

Taking the limit as $\Delta t \to 0$, we obtain the (linear) differential equation

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t). \tag{17.5.9}$$

The solution of Equation 17.5.9 is

$$P_0(t) = Ce^{-\lambda t}, \tag{17.5.10}$$

where C is an arbitrary constant. To evaluate C, we have the initial condition $P_0(0) = P[N(0)=0\,] = 1$ from Axion 1. Therefore,

$$P[N(t)=0\,] = P_0(t) = e^{-\lambda t}. \tag{17.5.11}$$

Next, let us find $P_1(t) = P[N(t)=1\,]$. We either have no arrivals in $[0, t-\Delta t]$ and one arrival in $(t-\Delta t, t]$ or one arrival in $[0, t-\Delta t]$ and no arrivals in $(t-\Delta t, t]$. These are the only two possibilities because there can be at most one arrival in a time interval Δt. The two events are mutually exclusive. Therefore,

$$\begin{aligned} P[N(t)=1\,] &= P[N(t-\Delta t)=0, N(t)-N(t-\Delta t)=1\,] \\ &\quad + P[N(t-\Delta t)=0, N(t)-N(t-\Delta t)=0\,] &\quad (17.5.12) \\ &= P[N(t-\Delta t)=0\,]P[N(t)-N(t-\Delta t)=1\,] \\ &\quad + P[N(t-\Delta t)=1\,]P[N(t)-N(t-\Delta t)=0\,] &\quad (17.5.13) \\ &= P[N(t-\Delta t)=0\,]P[N(t+\Delta t)-N(t)=1\,] \\ &\quad + P[N(t-\Delta t)=1\,]P[N(t+\Delta t)-N(t)=0\,]. &\quad (17.5.14) \end{aligned}$$

Equation 17.5.13 follows from independence while Equation 17.5.14 follows from stationarity. Introducing $P_1(t)$ in Equation 17.5.14 and using Axion 3,

$$P_1(t) = P_0(t - \Delta t)\lambda\Delta t + P_1(t - \Delta t)(1 - \lambda\Delta t), \tag{17.5.15}$$

or

$$\frac{P_1(t) - P_1(t - \Delta t)}{\Delta t} = -\lambda P_1(t - \Delta t) + \lambda P_0(t - \Delta t). \tag{17.5.16}$$

Taking the limit as $\Delta t \to 0$, we obtain

$$\frac{dP_1(t)}{dt} + \lambda P_1(t) = \lambda P_0(t). \tag{17.5.17}$$

In a similar manner, we can prove that

$$\frac{dP_k(t)}{dt} + \lambda P_k(t) = \lambda P_{k-1}(t), \tag{17.5.18}$$

where $k = 1, 2, 3, \ldots$ and $P_k(t) = P[N(t) = k]$.

This set of simultaneous linear equations can be solved recursively. Its solution is

$$P_k(t) = \exp(-\lambda t)\frac{(\lambda t)^k}{k!}, \qquad k = 0, 1, 2, \ldots, \tag{17.5.19}$$

which is the Poisson probability mass function. Here λ is the average number of arrivals per second.

In the realization of a Poisson process, one of the important quantities is the *arrival time*, t_n, shown in Figure 17.5.1. Of course, the arrival time is also a random process and will change with each new realization. A related quantity $Z_i = t_i - t_{i-1}$, the time intervals between two successive occurrences (interoccurrence times) of Poisson events. We will now show that the random variables Z_1, Z_2, etc., are independent and identically distributed with

$$P(Z_n \le x) = 1 - e^{-\lambda x}, \qquad x \ge 0, \quad n = 1, 2, 3, \ldots \tag{17.5.20}$$

We begin by noting that

$$P(Z_1 > t) = P[N(t) = 0] = e^{-\lambda t} \tag{17.5.21}$$

from Equation 17.5.19. Therefore, Z_1 has an exponential distribution.

Let us denote its probability density by $p_{Z_1}(z_1)$. From the joint conditional density function,

$$P(Z_2 > t) = \int_0^{\xi_1} P(Z_2 > t|Z_1 = z_1)p_{Z_1}(z_1)\, dz_1, \tag{17.5.22}$$

where $0 < \xi_1 < t$. If $Z_1 = z_1$, then $Z_2 > t$ if and only if $N(t + z_1) - N(z_1) = 0$. Therefore, using the independence and stationary properties,

$$P\{Z_2 > t|Z_1 = P[N(t + z_1) - N(z_1) = 0]\} = P[N(t) = 0] = e^{-\lambda t}. \tag{17.5.23}$$

Consequently,

$$P(Z_2 > t) = e^{-\lambda t}, \tag{17.5.24}$$

showing that Z_2 is also exponential. Also, Z_2 is independent of Z_1. Now, let us introduce $p_{Z_2}(z_2)$ as the probability density of $Z_1 + Z_2$. By similar arguments we can show that Z_3 is also exponential. The final result follows by induction.

• Example 17.5.1: Random telegraph signal

We can use the fact that interoccurrence times are independent and identically distributed to realize the Poisson process. An important application of this is in the generation of the random telegraph signal: $X(t) = (-1)^{N(t)}$. However, no one uses this definition to compute the signal; they use the arrival times to change the signal from $+1$ to -1 or vice versa.

We begin by noting that $T_i = T_{i-1} + Z_i$, with $i = 1, 2, \ldots$, $T_0 = 0$, and T_i is the ith arrival time. Each Z_i has the same exponent probability density function. From Equation 16.4.17,

$$Z_i = \frac{1}{\lambda} \ln\left(\frac{1}{1 - U_i}\right), \tag{17.5.25}$$

where the U_i's are from a uniform distribution. The realization of a random telegraphic signal is given by the MATLAB code:

```
clear
N = 100; % number of switches in realization
lambda = 0.15; % switching rate
X = [ ];
% generate N uniformly distributed random variables
S = rand(1,N);
% transform S into an exponential random variable
T = - log(S)/lambda;
V = cumsum(T); % compute switching times
t = [0.01:0.01:100]; % create time array
icount = 1; amplitude = -1; % initialize X(t)
for k = 1:10000
  if ( t(k) >= V(icount) ) % at each switching point
    icount = icount + 1;
    amplitude = - amplitude; % switch sign
  end
  X(k) = amplitude; % generate X(t)
end
plot(t,X) % plot results
xlabel('\it t','FontSize',25);
ylabel('\it X(t)/a','FontSize',25);
axis([0 max(t) -1.1 1.1])
```

This was the MATLAB code that was used to generate Figure 17.5.2. □

• Example 17.0.2

It takes a workman an average of one hour to put a widget together. Assuming that the task can be modeled as a Poisson process, what is the probability that a workman can build 12 widgets during an eight-hour shift?

The probability that n widgets can be constructed by time t is

$$P[N(t) = n] = e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \qquad (\mathbf{17.5.26})$$

Therefore, the probability that 12 or more widgets can be constructed in eight hours is

$$P[N(t) \geq 12] = e^{-8} \sum_{n=12}^{\infty} \frac{8^n}{n!} = 0.1119, \qquad (\mathbf{17.5.27})$$

since $\lambda = 1$.

We could have also obtained our results by creating 12 exponentially distributed time periods and summed them together using MATLAB:

```
t_uniform = rand(1,12);
T = - log(1-t_uniform);
total_time = sum(T);
```

Then, by executing this code a large number `N` of times and counting the number `icount` of times that `total_time <= 8`, the probability equals `icount / N`.

Problems

1. Use the generating function

$$F(z,t) = \sum_{n=0}^{\infty} p_n(t) z^n, \qquad |z| < 1,$$

with $F(z,0) = 1$ to solve Equation 17.5.18 by showing that $F(z,t) = e^{\lambda t(z-1)}$. Then, by expanding $F(z,t)$, recover Equation 17.5.19.

Project: Output from a Filter When the Input Is a Random Telegraphic Signal[11]

In the study of many systems, such as linear filters, the output $y(\cdot)$ can be written as

$$y(t) = \int_{-\infty}^{t} W(t-\tau) x(\tau)\, d\tau,$$

where $W(\cdot)$ is the weight function and $x(\cdot)$ is the input. The purpose of this project is to explore the probability density $P(y)$ of the output when $x(t)$ is the random telegraphic signal, a Poisson random process. You will filter this input two ways: (1) ideal integrator with finite memory: $W(t) = H(t) - H(t - \tau_1)$, $\tau_1 > 0$, and (2) simple $RC = 1$ low-pass filter $W(t) = e^{-t} H(t)$.

Step 1: Use MATLAB to code $x(t)$ where the expected time between the zeros is λ.

Step 2: Develop MATLAB code to compute $y(t)$ for each of the weight functions $W(t)$.

Step 3: Compute $P(y)$ for both filters. How do your results vary as λ varies?

[11] Suggested by a paper by McFadden, J. A., 1959: The probability density of the output of a filter when the input is a random telegraphic signal: Differential-equation approach. *IRE Trans. Circuit Theory*, **6**, 228–233.

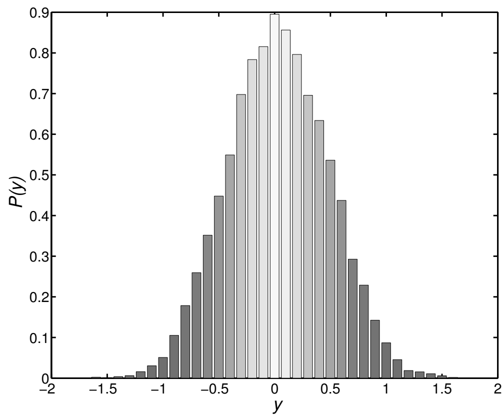


Figure 17.5.1: The probability density $P(y)$ of the output from an ideal integrator with finite memory when the input is a random telegraphic signal when $\Delta t = 0.01$, $\lambda = 2$, and $\tau_1 = 10$.

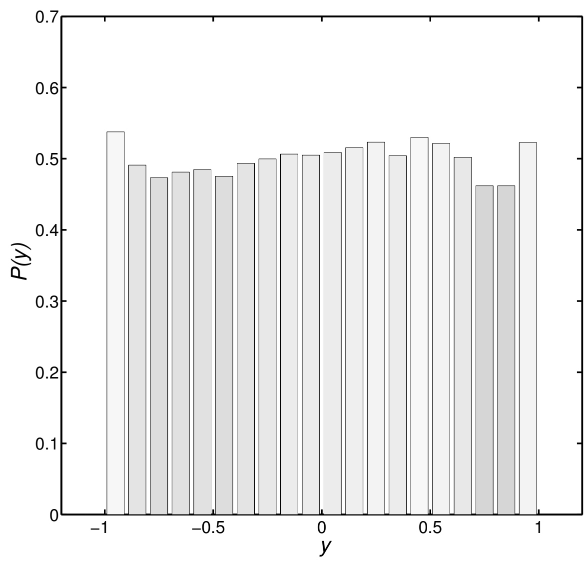


Figure 17.5.2: The probability density $P(y)$ of the output from a simple $RC = 1$ filter, when the input is a random telegraphic signal, when $\lambda = 1$, and $\Delta t = 0.05$.

Further Readings

Beckmann, P., 1967: *Probability in Communication Engineering.* Harcourt, Brace & World, 511 pp. A presentation of probability as it applies to problems in communication engineering.

Gillespie, D. T., 1991: *Markov Processes: An Introduction for Physical Scientists.* Academic Press, 592 pp. For the scientist who needs an introduction to the details of the subject.

Hsu, H., 1997: *Probability, Random Variables, & Random Processes.* McGraw-Hill, 306 pp.

Summary of results plus many worked problems.

Kay, S. M., 2006: *Intuitive Probability and Random Processes Using MATLAB*. Springer, 833 pp. A well-paced book designed for the electrical engineering crowd.

Ross, S. M., 2007: *Introduction to Probability Models.* Academic Press, 782 pp. An introductory undergraduate book in applied probability and stochastic processes.

Tuckwell, H. C., 1995: *Elementary Applications of Probability Theory.* Chapman & Hall, 292 pp. This book presents applications using probability theory, primarily from biology.

Chapter 18
Itô's Stochastic Calculus

In Chapter 1 we studied the solution to first-order ordinary differential equations

$$\frac{dx}{dt} = a(t, x), \qquad x(0) = x_0. \tag{18.0.1}$$

There we showed that Equation 18.0.1 has the solution

$$x(t) = x(0) + \int_0^t a[\eta, x(\eta)]\, d\eta. \tag{18.0.2}$$

Consider now the analogous *stochastic differential equation*:

$$dX(t) = a[t, X(t)]\, dt, \qquad X(0) = X_0. \tag{18.0.3}$$

Although Equations 18.0.1 and 18.0.3 formally appear the same, an immediate question is what is meant by $dX(t)$. In elementary calculus the concept of the infinitesimal involves limits, continuity, and so forth. As we shall see in Section 18.2, Brownian motion, a very common stochastic process, is nowhere differentiable. Here we can merely say that $dX(t) = X(t + dt) - X(t)$.

Consider now a modification of Equation 18.0.3 where we introduce a random forcing:

$$dX(t) = a[t, X(t)]\, dt + b[t, X(t)]\, dB(t), \qquad X(0) = X_0. \tag{18.0.4}$$

Here $dB(t) = B(t + dt) - B(t)$, $B(t)$ denotes Brownian motion and $a[t, X(t)]$ and $b[t, X(t)]$ are deterministic functions. Consequently, changes to $X(t)$ result from (1) the effects of the initial conditions and (2) noise generated by Brownian motion (the driving force). Stochastic processes governed by Equation 18.0.4 are referred to as *Itô processes*.

Following the methods used to derive 18.0.2 we can formally write the solution to Equation 18.0.4 as

$$X(t) = X_0 + \int_0^t a[\eta, X(\eta)]\, d\eta + \int_0^t b[\eta, X(\eta)]\, dB(\eta). \tag{18.0.5}$$

The first integral in Equation 18.0.5 is the conventional Riemann integral from elementary calculus and is well understood. The second integral, however, is new and must be treated with care. It is called *Itô's stochastic integral* and treated in Section 18.3.

In summary, a simple analog to first-order ordinary differential equations for a single random variable $X(t)$ raises several important questions. What is meant by the infinitismal and the integral in stochastic calculus? In this chapter we will focus on Itô processes and the associated calculus. Although Itô's calculus is an important discipline, it is not the only form of stochastic calculus. The interested student is referred elsewhere for further study.

Problems

1. The Poisson random process $N(t)$ is defined by

$$N(t) = \sum_{n=1}^{\infty} H(t - t_n),$$

where t_n is a sequence of independent and identically distributed inter-arrival times t_n. A graphical representation of $N(t)$ would consist of ever-increasing steps with the edges located at $t = t_n$. Use the definition of $dN(t) = N(t + dt) - N(t)$ to show that

$$dN(t) = \begin{cases} 1, & \text{for } t = t_n, \\ 0, & \text{otherwise.} \end{cases}$$

2. The telegraph signal is defined by $X(t) = (-1)^{N(t)}$, where $N(t)$ is given by the Poisson random distribution in Problem 1. Show[1] that

$$dX(t) = X(t + dt) - X(t) = (-1)^{N(t)} \left[(-1)^{dN(t)} - 1\right] = -2X(t)\, dN(t).$$

Hint: Consider $dN(t)$ at various times.

3. If $X(t)$ and $Y(t)$ denote two stochastic processes, use the definition of the derivative to show that (a) $d[cX(t)] = c\, dX(t)$, where c is a constant, (b) $d[X(t) \pm Y(t)] = dX(t) \pm dY(t)$, and (c) $d[X(t)Y(t)] = X(t)\, dY(t) + Y(t)\, dX(t) + dX(t)\, dY(t)$.

18.1 RANDOM DIFFERENTIAL EQUATIONS

A large portion of this book has been devoted to solving differential equations. Here we examine the response of differential equations to random forcing where the differential

[1] Taken from Janaswamy, R., 2013: On random time and on the relation between wave and telegraph equation. *IEEE Trans. Antennas Propag.*, **61**, 2735–2744.

equation describes a nonrandom process. This is an important question in the sciences and engineering because noise, a random phenomenon, is ubiquitous in nature.

Because the solution to random differential equations can be found by conventional techniques, we can use them to study the effect of randomness on the robustness of a solution to a differential equation subject to small changes of the initial condition. Although this may be of considerable engineering interest, it is really too simple to develop a deep understanding of stochastic differential equations.

• **Example 18.1.1: LR circuit**

One of the simplest differential equations that we encountered in Chapter 1 involves the mathematical model for an LR electrical circuit:

$$L\frac{dI}{dt} + RI = E(t), \tag{18.1.1}$$

where $I(t)$ denotes the current within an electrical circuit with inductance L and resistance R, and $E(t)$ is the mean electromotive force. If we solve this first-order ordinary differential equation using an integrating factor, its solution is

$$I(t) = I(0)\exp\left(-\frac{Rt}{L}\right) + \frac{1}{L}\exp\left(-\frac{Rt}{L}\right)\int_0^t F(\tau)\exp\left(\frac{R\tau}{L}\right)\,d\tau. \tag{18.1.2}$$

Clearly, if the electromotive forcing is random, so is the current.

In the previous chapter we showed that the mean and variance were useful parameters in characterizing a random variable. This will also be true here. If we find the mean of the solution,

$$E[I(t)] = I(0)\exp\left(-\frac{Rt}{L}\right) \tag{18.1.3}$$

provided $E[F(t)] = 0$. Thus, the mean of the current is the same as that for an ideal LR circuit.

Turning to the variance,

$$\begin{aligned}\sigma_I^2(t) &= E[I^2(t)] - \{E[I(t)]\}^2 && (18.1.4)\\ &= E\left[I^2(0)\exp\left(-\frac{2Rt}{L}\right)\right]\\ &\quad + \frac{2I(0)}{L}\exp\left(-\frac{2Rt}{L}\right)\int_0^t E[F(\tau)]\exp\left(\frac{R\tau}{L}\right)\,d\tau\\ &\quad + \frac{1}{L^2}\exp\left(-\frac{2Rt}{L}\right)\int_0^t\int_0^t E[F(\tau)F(\tau')]\exp\left[\frac{R(\tau+\tau')}{L}\right]\,d\tau'\,d\tau && (18.1.5)\\ &\quad - I^2(0)\exp\left(-\frac{2Rt}{L}\right)\\ &= \frac{1}{L^2}\exp\left(-\frac{2Rt}{L}\right)\int_0^t\int_0^t E[F(\tau)F(\tau')]\exp\left[\frac{R(\tau+\xi)}{L}\right]\,d\tau'\,d\tau. && (18.1.6)\end{aligned}$$

To proceed further we need the autocorrelation $E[F(\tau)F(\tau')]$. In papers by Ornstein et al.[2] and Jones and McCombie,[3] they adopted a random process with the autocorrelation

---

[2] Ornstein, L. S., H. C. Burger, J. Taylor, and W. Clarkson, 1927: The Brownian movement of a galvanometer and the influence of the temperature of the outer circuit. *Proc. Roy. Soc. London, Ser. A*, **115**, 391–406.

[3] Jones, R. V., and C. W. McCombie, 1952: Brownian fluctuations in galvanometer and galvanometer amplifiers. *Phil. Trans. Roy. Soc. London, Ser. A*, **244**, 205–230.

function

$$E[F(\tau)F(\tau')] = 2D\delta(\tau - \tau'). \tag{18.1.7}$$

The advantage of this process is that it is mathematically the simplest because it possesses a white power spectrum. Unfortunately this random process can never be physically realized because it would possess infinite mean square power. All physically realizable processes involve a power spectrum that tends to zero at sufficiently high frequencies. If $\Phi(\omega)$ denotes the power spectrum, this condition can be expressed as

$$\int_0^\infty \Phi(\omega)\, d\omega < \infty. \tag{18.1.8}$$

In view of these considerations, let us adopt the autocorrelation

$$R_X(\tau - \tau') = \int_0^\infty \Phi(\omega) \cos[\omega(\tau - \tau')]\, d\omega, \tag{18.1.9}$$

where $\Phi(\omega)$ is the power spectrum of $F(\tau)$. Therefore, the variance becomes

$$\sigma_I^2(t) = \frac{1}{L^2} \int_0^t \int_0^t \int_0^\infty \Phi(\omega) \exp\left[-\frac{R(t-\tau)}{L}\right] \exp\left[-\frac{R(t-\tau')}{L}\right] \cos[\omega(\tau - \tau')]\, d\omega\, d\tau\, d\tau'. \tag{18.1.10}$$

Reversing the ordering of integration,

$$\sigma_I^2(t) = \frac{1}{L^2} \int_0^\infty \Phi(\omega) \int_0^t \int_0^t \exp\left[-\frac{R(2t - \tau - \tau')}{L}\right] \cos[\omega(\tau - \tau')]\, d\tau\, d\tau'\, d\omega. \tag{18.1.11}$$

We can evaluate the integrals involving τ and τ' exactly. Equation 18.1.11 then becomes

$$\sigma_I^2(t) = \int_0^\infty \frac{\Phi(\omega)}{\omega^2 + R^2/L^2} \left[1 + e^{-2Rt/L} - 2e^{-Rt/L} \cos(\omega t)\right] d\omega. \tag{18.1.12}$$

Let us now consider some special cases. As $t \to 0$, $\sigma_I^2(t) \to 0$ and the variance is initially small. On the other hand, as $t \to \infty$,

$$\sigma_I^2(t) = \int_0^\infty \frac{\Phi(\omega)}{\omega^2 + R^2/L^2}\, d\omega. \tag{18.1.13}$$

Thus, the variance grows to a constant value, which we would have found by using Fourier transforms to solve the differential equation.

Consider now the special case $\Phi(\omega) = 2D/\pi$, a forcing by white noise. Ignoring the defects in this model, we can evaluate the integrals in Equation 18.1.13 exactly and find that

$$\sigma_I^2(t) = \frac{DL}{R} \left(1 - e^{-2Rt/L}\right). \tag{18.1.14}$$

These results are identical to those found by Uhlenbeck and Ornstein[4] in their study of a free particle in Brownian motion. □

---

[4] Uhlenbeck, G. E., and L. S. Ornstein, 1930: On the theory of the Brownian motion. *Phys. Review*, **36**, 823–841. See the top of their page 828.

• **Example 18.1.2: Damped harmonic motion**

Another classic differential equation that we can excite with a random process is the damped harmonic oscillator:

$$y'' + 2\xi\omega_0 y' + \omega_0^2 y = F(t), \tag{18.1.15}$$

where $0 \le \xi < 1$, y denotes the displacement, t is time, $\omega_0^2 = k/m$, $2\xi\omega_0 = \beta/m$, m is the mass of the oscillator, k is the linear spring constant, and β denotes the constant of a viscous damper. In Chapter 2 we showed how to solve this second-order ordinary differential equation. Its solution is

$$y(t) = y(0)e^{-\xi\omega_0 t}\left[\cos(\omega_1 t) + \frac{\xi\omega_0}{\omega_1}\sin(\omega_1 t)\right] + \frac{y'(0)}{\omega_1}e^{-\xi\omega t}\sin(\omega_1 t) + \int_0^t h(t-\tau)F(\tau)\,d\tau, \tag{18.1.16}$$

where $\omega_1 = \omega_0\sqrt{1-\xi^2}$, and

$$h(t) = \frac{e^{-\xi\omega_0 t}}{\omega_1}\sin(\omega_1 t)H(t). \tag{18.1.17}$$

Again we begin by finding the mean of Equation 18.1.16. It is

$$E[y(t)] = y(0)e^{-\xi\omega_0 t}\left[\cos(\omega_1 t) + \frac{\xi\omega_0}{\omega_1}\sin(\omega_1 t)\right] + \frac{y'(0)}{\omega_1}e^{-\xi\omega t}\sin(\omega_1 t) + \int_0^t h(t-\tau)E[F(\tau)]\,d\tau. \tag{18.1.18}$$

If we again choose a random process where $E[F(t)] = 0$, the integral vanishes and the stochastic mean of the motion only depends on the initial conditions.

Turning to the variance,

$$\sigma_I^2(t) = E[y^2(t)] - \{E[y(t)]\}^2 = \int_0^t \int_0^t h(t-\tau)h(t-\tau')E[F(\tau)F(\tau')]\,d\tau\,d\tau'. \tag{18.1.19}$$

If we again adopt the autocorrelation function

$$R_X(\tau - \tau') = \int_0^\infty \Phi(\omega)\cos[\omega(\tau-\tau')]\,d\omega, \tag{18.1.20}$$

where $\Phi(\omega)$ is the power spectrum of $F(\tau)$, then

$$\sigma_I^2(t) = \int_0^\infty \frac{\Phi(\omega)}{\omega_1^2}\int_0^t \int_0^t e^{-\xi\omega_0(2t-\tau-\tau')}\sin[\omega_1(t-\tau)]\sin[\omega_1(t-\tau')]\cos[\omega(\tau-\tau')]\,d\tau\,d\tau'\,d\omega. \tag{18.1.21}$$

Carrying out the integrations in τ and τ', we finally obtain

$$\begin{aligned}\sigma_I^2(t) = \int_0^\infty \frac{\Phi(\omega)}{|\Omega(\omega)|^2}\Bigg(1 + e^{-2\xi\omega_0 t}\Bigg\{1 + \frac{2\xi\omega_0}{\omega_1}\sin(\omega_1 t)\cos(\omega_1 t)\\ - e^{\omega_0\xi t}\left[2\cos(\omega_1 t) + \frac{2\xi\omega_0}{\omega_1}\sin(\omega_1 t)\right]\cos(\omega t) - e^{\xi\omega_0 t}\frac{2\omega}{\omega_1}\sin(\omega_1 t)\sin(\omega t)\\ + \frac{\xi^2\omega_0^2 - \omega_1^2 + \omega^2}{\omega_1^2}\sin^2(\omega_1 t)\Bigg\}\Bigg)\,d\omega,\end{aligned} \tag{18.1.22}$$

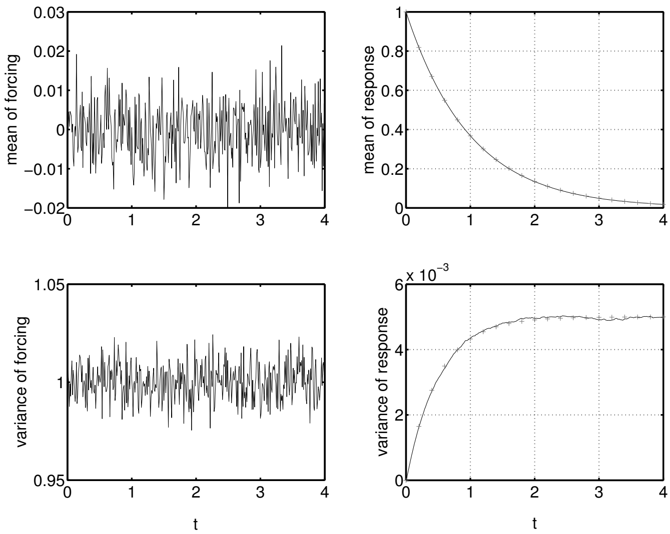


Figure 18.1.1: The mean and variance of the response for the differential equation $y' + y = f(t)$ when forced by Gaussian random noise. The parameters used are $y(0) = 1$ and $\Delta\tau = 0.01$.

where $|\Omega(\omega)|^2 = (\omega_0^2 - \omega^2)^2 + 4\omega^2\omega_0^2\xi^2$.

As in the previous example, $\sigma_I^2(t) \to 0$ as $t \to 0$ and the variance is initially small. The steady-state variance now becomes

$$\sigma_I^2(t) = \int_0^\infty \frac{\Phi(\omega)}{|\Omega(\omega)|^2}\, d\omega. \tag{18.1.23}$$

Finally, for the special case $\Phi(\omega) = 2D/\pi$, the variance is

$$\sigma_I^2(t) = \frac{D}{2\xi\omega_0^2}\left\{1 - \frac{e^{-2\xi\omega_0 t}}{\omega_1^2}\left[\omega_1^2 + \omega_0\omega_1\xi\sin(2\omega_1 t) + 2\xi^2\omega_0^2\sin^2(\omega_1 t)\right]\right\}. \tag{18.1.24}$$

These results are identical to those found by Uhlenbeck and Ornstein[5] in their study of a harmonically bound particle in Brownian motion.

Project: Low-Pass Filter with Random Input

Consider the initial-value problem

$$y' + y = f(t), \qquad y(0) = y_0.$$

It has the solution

$$y(t) = y_0 e^{-t} + e^{-t}\int_0^t e^\tau f(\tau)\, d\tau.$$

This differential equation is identical to that governing an RC electrical circuit. This circuit has the property that it filters out high frequency disturbances. Here we explore the case when $f(t)$ is a random process.

[5] Ibid. See their pages 834 and 835.

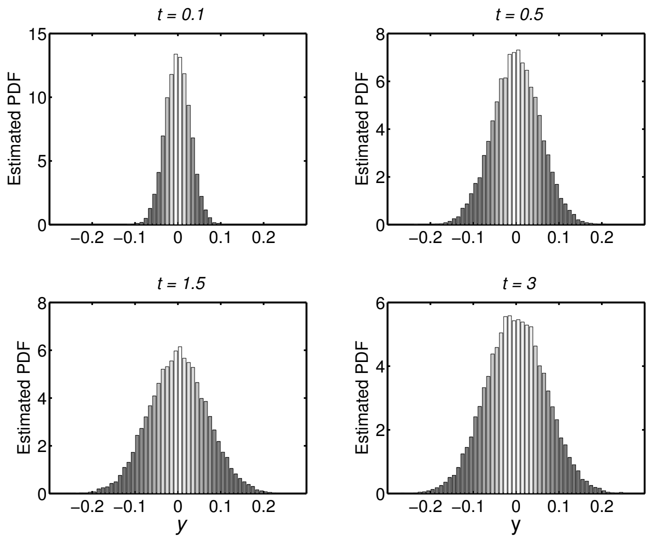


Figure 18.1.2: The probability density function for the response to the differential equation $y' + y = f(t)$ when $f(t)$ is a Gaussian distribution. Twenty thousand realizations were used to compute the density function. Here the parameters used are $y(0) = 0$ and $\Delta\tau = 0.01$.

Step 1: Using the MATLAB intrinsic function `randn`, generate a stationary white noise excitation of length `N`. Let `deltat` denote the time interval Δt between each new forcing so that `n = 1` corresponds to `t = 0` and `n = N` corresponds to the end of the record `t = T`.

Step 2: Using the Gaussian random forcing that you created in Step 1, develop a MATLAB code to compute $y(t)$ given $y(0)$ and $f(t)$.

Step 3: Once you have confidence in your code, modify it so that you can generate many realizations of $y(t)$. Save your solution as a function of t and realization. Use MATLAB's intrinsic functions `mean` and `var` to compute the mean and variance as a function of time. Figure 18.1.1 shows the results when 2000 realizations were used. For comparison the mean and variance of the forcing have also been included. Ideally this mean and variance should be zero and one, respectively. We have also included the exact mean and variance, given by Equation 18.1.3 and Equation 18.1.14, when we set $L = R = 1$ and $D = \Delta t/2$.

Step 4: Now generalize your MATLAB code so that you can compute the probability density function of finding $y(t)$ lying between y and $y + dy$ at various times. Figure 18.1.2 illustrates four times when $y(0) = 0$ and $\Delta\tau = 0.01$.

Step 5: Modify your MATLAB code so that you can compute the autocovariance. See Figure 18.1.3.

Project: First-Passage Problem with Random Vibrations[6]

In the design of devices, it is often important to know the chance that the device will exceed its design criteria. In this project you will examine how often the amplitude of a

[6] Based on a paper by Crandall, S. H., K. L. Chandiramani, and R. G. Cook, 1966: Some first-passage problems in random vibration. *J. Appl. Mech.*, **33**, 532–538.

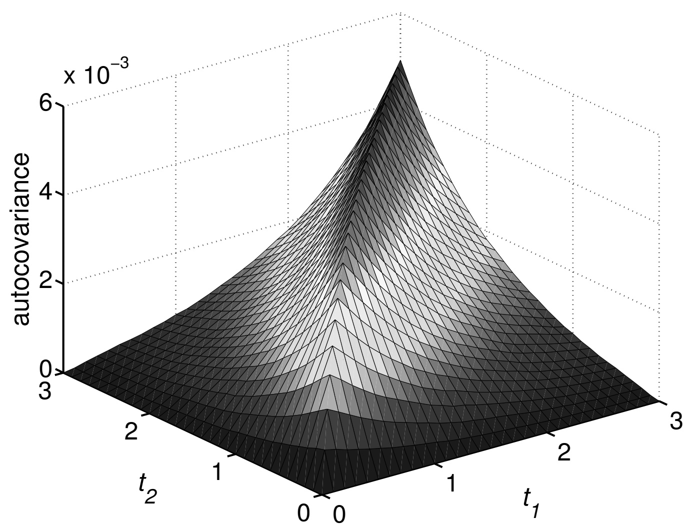


Figure 18.1.3: The autocovariance function for the differential equation $y' + y = f(t)$ when $f(t)$ is a Gaussian distribution. Twenty thousand realizations were used. The parameters used here are $y(0) = 0$ and $\Delta\tau = 0.01$.

simple, slightly damped harmonic oscillator

$$y'' + 2\zeta\omega_0 y' + \omega_0^2 y = f(t), \qquad 0 < \zeta \ll 1, \tag{1}$$

will exceed a certain magnitude when forced by white noise. In the physical world this transcending of a barrier or passage level leads to "bottoming" or "short circuiting."

Step 1: Using the MATLAB command `randn`, generate a stationary white noise excitation of length `N`. Let `deltat` denote the time interval Δt between each new forcing so that `n = 1` corresponds to `t = 0` and `n = N` corresponds to the end of the record `t = T`.

Step 2: The exact solution to Equation (1) is

$$\begin{aligned} y(t) &= y(0)e^{-\zeta\omega_0 t}\left[\cos\left(\sqrt{1-\zeta^2}\,\omega_0 t\right) + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t\right)\right] \\ &\quad + \frac{y'(0)}{\omega_0\sqrt{1-\zeta^2}}e^{-\zeta\omega_0 t}\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t\right) \qquad (2) \\ &\quad + \int_0^t e^{-\zeta\omega_0(t-\tau)}\frac{\sin\left[\sqrt{1-\zeta^2}\,\omega_0(t-\tau)\right]}{\sqrt{1-\zeta^2}}\frac{f(\tau)}{\omega_0^2}\,d(\omega_0\tau) \\ &= y(0)e^{-\zeta\omega_0 t}\left[\cos\left(\sqrt{1-\zeta^2}\,\omega_0 t\right) + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t\right)\right] \\ &\quad + \frac{y'(0)}{\omega_0\sqrt{1-\zeta^2}}e^{-\zeta\omega_0 t}\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t\right) \qquad (3) \\ &\quad + e^{-\zeta\omega_0 t}\frac{\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t\right)}{\sqrt{1-\zeta^2}}\int_0^t e^{\zeta\omega_0\tau}\cos\left(\sqrt{1-\zeta^2}\,\omega_0\tau\right)\frac{f(\tau)}{\omega_0^2}\,d(\omega_0\tau) \end{aligned}$$

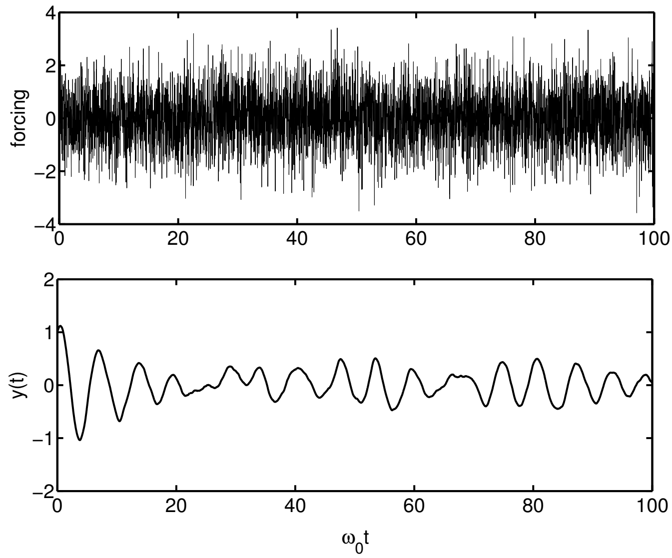


Figure 18.1.4: A realization of the random function $y(t)$ governed by Equation (1) when forced by the Gaussian random forcing shown in the top frame. The parameters used here are $y(0) = 1$, $y'(0) = 0.5$, $\zeta = 0.1$, and $\omega_0\Delta\tau = 0.02$.

$$- e^{-\zeta\omega_0 t}\frac{\cos\left(\sqrt{1-\zeta^2}\,\omega_0 t\right)}{\sqrt{1-\zeta^2}}\int_0^t e^{\zeta\omega_0\tau}\sin\left(\sqrt{1-\zeta^2}\,\omega_0\tau\right)\frac{f(\tau)}{\omega_0^2}\,d(\omega_0\tau).$$

Because you will be computing numerous realizations of $y(t)$ for different $f(t)$'s, an efficient method for evaluating the integrals must be employed. Equation (3) is more efficient than Equation (2).

Using the Gaussian random forcing that you created in Step 1, develop a MATLAB code to compute $y(t)$ given $y(0)$, $y'(0)$, ζ and $f(t)$. Figure 18.1.4 illustrates a realization where the trapezoidal rule was used to evaluate the integrals in Equation (3).

Step 3: Now that you can compute $y(t)$ or `y(n)` for a given Gaussian random forcing, generalize your code so that you can compute `irun` realizations and store them in `y(n,m)` where `m = 1:irun`. For a specific `n` or $\omega_0 t$, you can use MATLAB's commands `mean` and `var` to compute the mean $\mu_X(t)$ and the variance $\sigma_X^2(t)$. Figure 18.1.5 shows the results when 1000 realizations were used. For comparison the mean and variance of the forcing have also been included. Ideally this mean and variance should be zero and one, respectively. The crosses give the exact results that

$$\mu_X(t) = y(0)e^{-\zeta\omega_0 t}\left[\cos\left(\sqrt{1-\zeta^2}\,\omega_0 t\right) + \frac{\zeta}{\sqrt{1-\zeta^2}}\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t\right)\right]$$
$$+ \frac{y'(0)}{\omega_0\sqrt{1-\zeta^2}}e^{-\zeta\omega_0 t}\sin\left(\sqrt{1-\zeta^2}\,\omega_0 t\right) \qquad (4)$$

and Equation 18.1.25 when $D = \omega_0\Delta t/2$.

Step 4: Finally generalize your MATLAB code so that you store the time `T(m)` that the solution `y(n)` exceeds a certain amplitude $b > 0$ for the *first* time during the realization `m`.

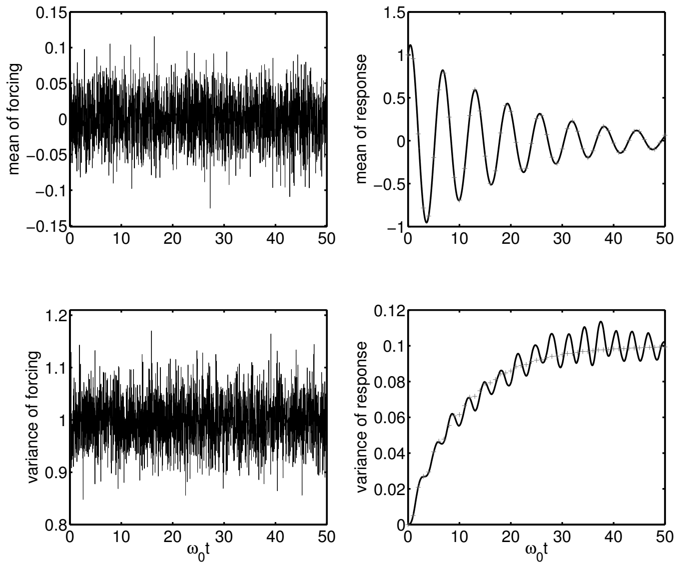


Figure 18.1.5: The mean $\mu_X(t)$ and variance $\sigma_X^2(t)$ of a slightly damped simple harmonic oscillator when forced by the Gaussian random noise. The parameters used here are $y(0) = 1$, $y'(0) = 0.5$, $\zeta = 0.1$, and $\omega_0 \Delta\tau = 0.02$.

Of course, you can do this for several different b's during a particular realization. Once you have this data you can estimate the probability density function using `histc`. Figure 18.1.6 illustrates four probability density functions for $b = 0.4$, $b = 0.8$, $b = 1.2$, and $b = 1.6$.

Project: Wave Motion Generated by Random Forcing[7]

In the previous projects we examined ordinary differential equations that we forced with a random process. Here we wish to extend our investigation to the one-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = \cos(\omega t)\delta[x - X(t)],$$

subject to the boundary conditions

$$\lim_{|x|\to\infty} u(x,t) \to 0, \qquad 0 < t,$$

and initial conditions

$$u(x,0) = u_t(x,0) = 0, \qquad -\infty < x < \infty.$$

Here ω is a constant and $X(t)$ is a stochastic process.

In Example 15.4.4 we show that the solution to this problem is

$$u(x,t) = \tfrac{1}{2}\int_0^t H[t - \tau - |X(\tau) - x|] \cos(\omega\tau)\, d\tau.$$

[7] Based on a paper by Knowles, J. K., 1968: Propagation of one-dimensional waves from a source in random motion. *J. Acoust. Soc. Am.*, **43**, 948–957.

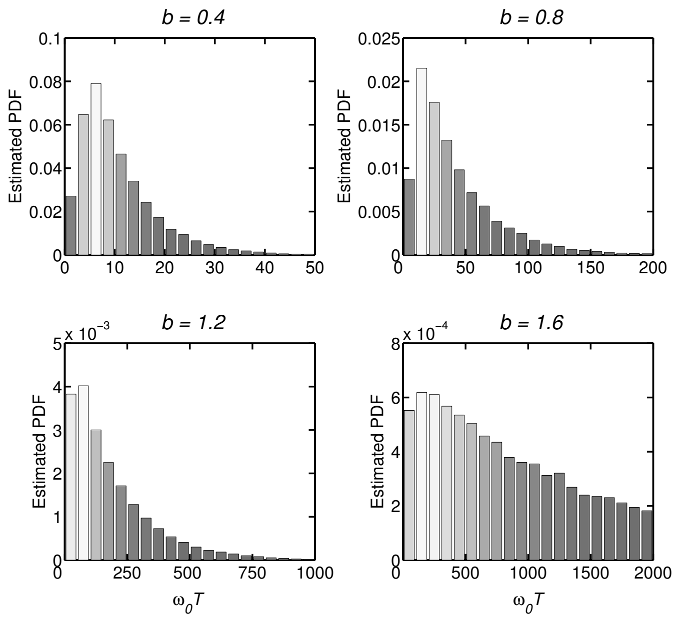


Figure 18.1.6: The probability density function that a slightly damped oscillator exceeds b at the time $\omega_0 T$. Fifty thousand realizations were used to compute the density function. The parameters used here are $y(0) = 0$, $y'(0) = 0$, $\zeta = 0.05$, and $\omega_0 \Delta\tau = 0.05$. The mean value of $\omega_0 T$ is 10.7 when $b = 0.4$, 41.93 when $b = 0.8$, 188.19 when $b = 1.2$, and 1406.8 when $b = 1.6$.

When the stochastic forcing is absent $X(t) = 0$, we can evaluate the integral and find that

$$u(x,t) = \frac{1}{2\omega} H(t - |x|) \sin[\omega(t - |x|)].$$

Step 1: Invoking the MATLAB command `randn`, use this Gaussian distribution to numerically generate an excitation $X(t)$.

Step 2: Using the Gaussian distribution from Step 1, develop a MATLAB code to compute $u(x,t)$. Figure 18.1.7 illustrates one realization where the trapezoidal rule was used to evaluate the integral.

Step 3: Now that you can compute $u(x,t)$ for a particular random forcing, generalize your code so that you can compute `irun` realizations. Then, for particular values of x and t, you can compute the corresponding mean and variance from the `irun` realizations. Figure 18.1.8 shows the results when 10,000 realizations were used.

Step 4: Redo your calculations but use a sine wave with random phase: $X(t) = A\sin(\Omega t + \xi)$, where A and Ω are constants and ξ is a random variable with a uniform distribution on $[0, 2\pi]$.

18.2: RANDOM WALK AND BROWNIAN MOTION

In 1827 the Scottish botanist Robert Brown (1773–1858) investigated the fertilization process in a newly discovered species of flower. Brown observed under the microscope that when the pollen grains from the flower were suspended in water, they performed a "rapid

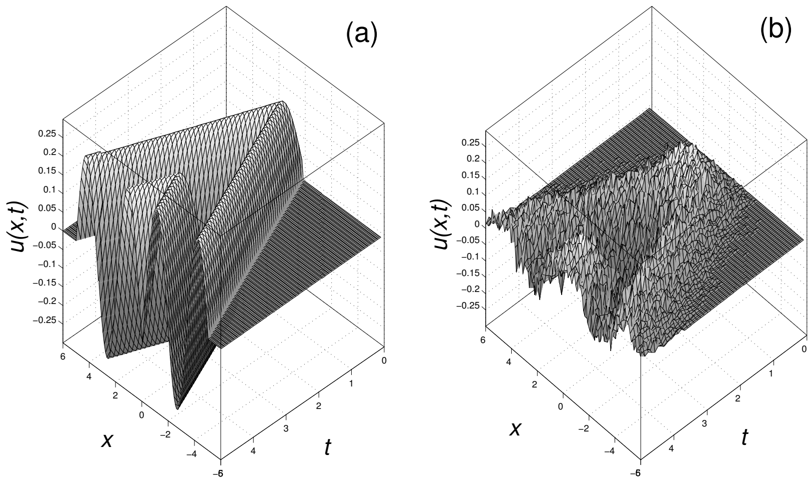


Figure 18.1.7: The solution (realization) of the wave equation when forced by a Gaussian distribution and $\omega = 2$. In frame (a), there is no stochastic forcing $X(t) = 0$. Frame (b) shows one realization.

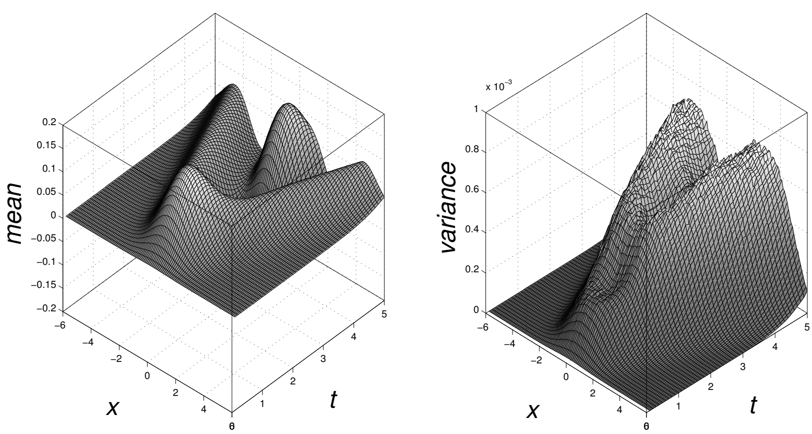


Figure 18.1.8: The mean and variance when the wave equation is forced by the stochastic forcing $\cos(\omega t)\delta[x - X(t)]$, where $\omega = 2$ and $X(t)$ is a Gaussian distribution.

oscillation motion." This motion, now known as *Brownian motion*, results from the random kinetic strikes on the pollen grain by water molecules. Brownian motion is an example of a random process known as *random walk*. This process has now been discovered in such diverse disciplines, from biology[8] to finance. In this section we examine its nature.

Consider a particle that moves along a straight line in a series of steps of equal length. Each step is taken, either forwards or backwards, with equal probability $\frac{1}{2}$. After taking N steps the particle *could* be at any one (let us denote it m) of the following points:

[8] Codling, E. A., M. J. Plank, and S. Benhamou, 2008: Random walk models in biology. *J. R. Soc. Interface*, **5**, 813–834.

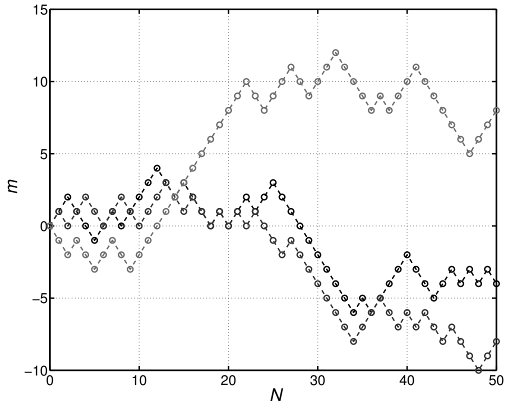


Figure 18.2.1: Three realizations of a one-dimensional random walk where $N = 50$.

$-N, -N+1, \ldots, -1, 0, 1, \ldots, N-1$ and N. Here m is a random variable.

We can generate realizations of one-dimensional Brownian motion using the MATLAB code:

```
clear

NN = 50; % select the number of steps for the particle to take
t = (0:1:NN); % create ''time'' as the particle moves

% create an array to give the position at each time step
m = zeros(size(t));
m(1) = 0; % initialize the position of particle

for N = 1:NN % now move the particle
  x = rand(1); % generate a random variable lying between [0,1]
  if (x <= 0.5) step = 1; % if less then 0.5, make it a ''head''
  else step = -1; end % otherwise it is a ''tail''
% move the particle one step to the right or left
  m(N+1) = m(N) + step;
end

%     plot the results

hold on
plot(t,m,'--ko','LineWidth',2,'MarkerSize',8)
xlabel('N','FontSize',25); ylabel('m','FontSize',25)
grid on % add a grid to axes
```

Figure 18.2.1 illustrates three such realizations.

A natural question would now be: What are the quantitative properties of random walk? In particular, what is the probability $P(m, N)$ that the particle is at point m after N displacements? We begin by noting the probability of any given sequence of N steps is $\left(\frac{1}{2}\right)^N$. The desired probability $P(m, N)$ equals $\left(\frac{1}{2}\right)^N$ times the number of distinct sequences of steps that will lead to the point m after N steps. To reach m, we must take $(N+m)/2$ steps in the positive direction and $(N-m)/2$ in the negative direction since $(N+m)/2-(N-m)/2 = m$. (Note both m *and* N must be even or odd.) The number of these distinct sequences is

$$\frac{N!}{\left[\frac{1}{2}(N+m)\right]!\left[\frac{1}{2}(N-m)\right]!}. \tag{18.2.1}$$

Therefore,

$$P(m,N) = \frac{N!}{\left[\frac{1}{2}(N+m)\right]!\left[\frac{1}{2}(N-m)\right]!}\left(\frac{1}{2}\right)^N. \tag{18.2.2}$$

Comparing these results with Equation 16.6.14, we see that $P(m,N)$ is simply a binomial distribution. For this reason, we immediately know $E(m) = 0$ and $\text{Var}(m) = N$. That is, the *average* position is the origin and the *spread* of the Brownian motion occurs as the square root of steps taken increases.

The case of greatest interest arises when N is large and $m \ll N$. Then we can approximate $P(m,N)$ by the Poisson distribution,

$$P(m,N) \approx \sqrt{\frac{2}{\pi N}}\exp\left(-\frac{m^2}{2N}\right). \tag{18.2.3}$$

Let us reexpress Equation 18.2.3 in terms of x and t where $x = m\Delta x$ and $t = N\Delta t$. Using these definitions, our equation becomes

$$P(x,t) = \frac{1}{2\sqrt{\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right), \tag{18.2.4}$$

if we define $D = (\Delta x)^2/(2\Delta t)$. The attentive student will note that $P(x,t)$ is the Green's function for the heat function, Example 15.5.1.

An alternative approach to this problem would be to compute many random walks and then calculate the probability density function from these computations. We can construct a MATLAB code to do this. First we would realize many random walks (here 2000) and count the number of times that they end at position m:

```
clear

NN = 100; % set the end point of the random walks
% introduce intermediate positions along the random walk
t = (0:1:NN);
% initialize array ``m'' which gives the position at any time
m = zeros(size(t));

for icount = 1:2000 % now perform many random walks

  m(1) = 0; % initial position of particle in each walk

  for N = 1:NN
    x = rand(1); % create a random variable between [0,1]
% if ''x'' less than 0.5, we have a ''heads''
    if (x <= 0.5) step = 1;
    else step = -1; end % otherwise we have a ''tail''
```

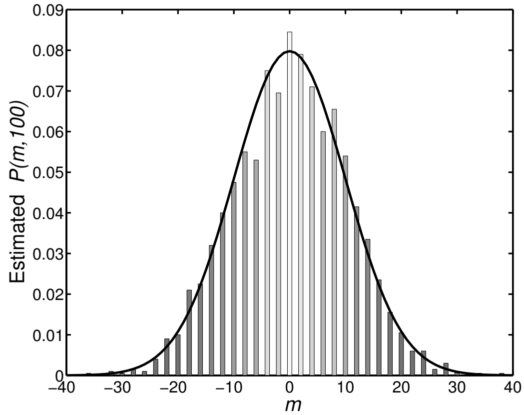


Figure 18.2.2: Numerical computation of $P(m, 100)$ using 2000 random walks. The black line gives Equation 18.2.3.

```
    m(N+1) = m(N) + step; % now take a step forward or backward
  end

% set up array that tracks of the final position of the particle
  location(icount) = m(N+1);

end

xx = -40:1:40;
% now count the particles that ended somewhere
%     between -40 and 40
[n,xout] = hist(location,xx)
% for comparison, compute Equation 18.2.3
w_exact = sqrt(2/(pi*NN))*exp(-xout.*xout/(2*NN));
n = n / 2000; % now compute the mass probability function

%     plot the results
bar_h = bar(xout,n)
bar_child = get(bar_h,'Children')
set(bar_child,'CData',n)
colormap(Autumn)
hold on
plot(xout,w_exact,'-k','LineWidth',3)
xlabel('\it m','FontSize',25)
ylabel('Estimated \it P(m,100)','FontSize',25)
```

Figure 18.2.2 illustrates the results of simulating random walk.

• Example 18.2.1: On the probability of striking a barrier

An important question in engineering is what is the probability that a given random system will exceed its design constraints. Here we ask a similar question about one-dimensional

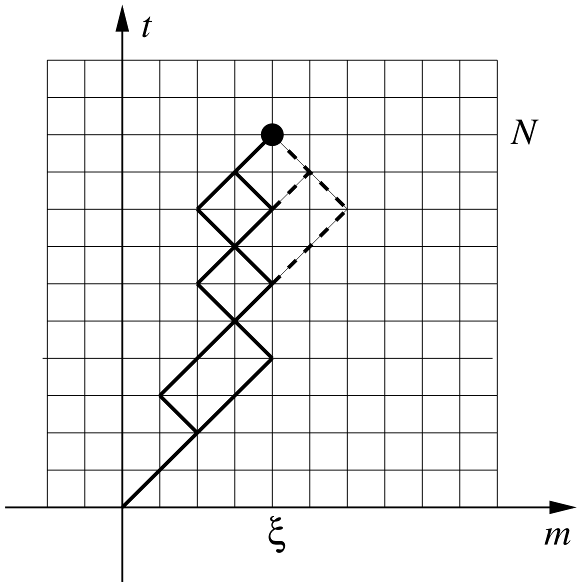


Figure 18.2.3: Several random walks from the origin to point (ξ, N). All of these walks would be excluded from our calculations because they either cross or touch the line $m = \xi$ before the final step.

Brownian motion: What is the probability that after taking N steps the particle arrives at ξ without ever having touched or crossed the line $m = \xi$ at *any* earlier step? We will do it exactly and then confirm our results using MATLAB.

The arrival of the particle at ξ after N steps implies that its position after $N-1$ steps must have been either $\xi - 1$ or $\xi + 1$. However, a trajectory from $(\xi+1, N-1)$ to (ξ, N) is not allowed because it must have crossed the line $m = \xi$ earlier. On the other hand, not all trajectories arriving at (ξ, N) from $(\xi-1, N-1)$ are acceptable because a certain number will have touched or crossed the line $m = \xi$ earlier than its last step. See Figure 18.2.3. Thus the number of permitted ways of arriving at ξ for the first time after N steps equals all possible ways of arriving at ξ *minus* any arrivals from $(\xi-1, N-1)$ *and* any arrivals that crossed or touched the line $m = \xi$ earlier than the $N-1$.

From our previous work, the number of possible ways from the origin to (ξ, N) is

$$\frac{N!}{\left[\frac{1}{2}(N+\xi)\right]!\left[\frac{1}{2}(N-\xi)\right]!}. \tag{18.2.5}$$

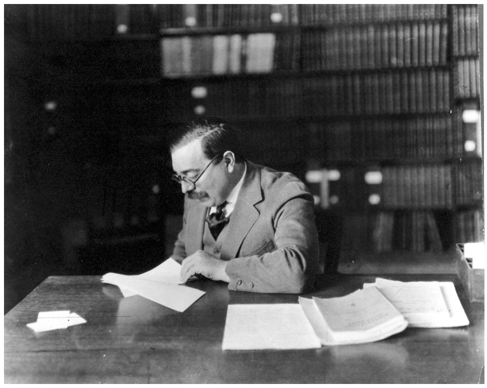
The number of possible ways from the origin to $(\xi+1, N-1)$ is

$$\frac{(N-1)!}{\left[\frac{1}{2}(N+\xi)\right]!\left[\frac{1}{2}(N-\xi-2)\right]!}. \tag{18.2.6}$$

Finally, the number of trajectories arriving at $(\xi-1, N-1)$ but having an earlier contact with, or a crossing of, the line $m = \xi$ is also

$$\frac{(N-1)!}{\left[\frac{1}{2}(N+\xi)\right]!\left[\frac{1}{2}(N-\xi-2)\right]!}, \tag{18.2.7}$$

since it equals the number of trajectories that arrive at $(\xi+1, N-1)$. From Figure 18.2.3 we see that, due to symmetry, the trajectory that leads to $(\xi+1, N-1)$ also leads to



One of the great mathematicians of the twentieth century, Norbert Wiener (1894–1964) graduated from high school at the age of 11 and Tufts at 14. Obtaining a doctorate in mathematical logic at 18, he repeatedly traveled to Europe for further education. His work extends over an extremely wide range from stochastic processes to harmonic analysis to cybernetics. (Photo courtesy of the MIT Museum with permission.)

$(\xi-1, N-1)$. Consequently the number of trajectories from the origin to (ξ, N) that have never touched or crossed $m=\xi$ is

$$\frac{N!}{\left[\frac{1}{2}(N+\xi)\right]!\left[\frac{1}{2}(N-\xi)\right]!}-2\frac{(N-1)!}{\left[\frac{1}{2}(N+\xi)\right]!\left[\frac{1}{2}(N-\xi-2)\right]!}, \tag{18.2.8}$$

or

$$\frac{\xi}{N}\frac{N!}{\left[\frac{1}{2}(N+\xi)\right]!\left[\frac{1}{2}(N-\xi)\right]!}. \tag{18.2.9}$$

The probability $P(\xi, N)$ that we are seeking is

$$P(\xi, N)=\frac{\xi}{N}\frac{N!}{\left[\frac{1}{2}(N+\xi)\right]!\left[\frac{1}{2}(N-\xi)\right]!}\left(\frac{1}{2}\right)^N. \tag{18.2.10}$$

For large N, $P(\xi, N)$ is approximately given by

$$P(\xi, N)=\frac{\xi}{N}\sqrt{\frac{2}{\pi N}}\exp\left(-\frac{\xi^2}{2N}\right). \tag{18.2.11}$$

We can also compute this probability using the MATLAB code given above. In this code we replace the counting process `location(icount) = m(N+1);` by

```
b = sort(m);

if ( (m(NN+1) == b(NN+1)) & (b(NN+1)>b(NN)) )
jcount = jcount + 1;
location(jcount) = m(NN+1);
end
```

where we initialize `jcount = 0` at the beginning. The idea behind this code is as follows: For each of the `icount` trajectories, we use the MATLAB routine `sort` to arrange them from

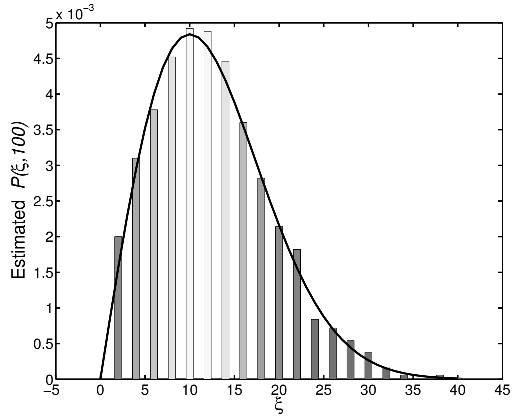


Figure 18.2.4: The probability $P(\xi, N)$ that a particle will reach the point $m = \xi$ without the particle ever crossing or touching the line $m = \xi$ earlier than $N = 100$. The solid line is the theoretical probability given by Equation 18.2.11. Here 50,000 random walks were taken.

the left-most to the right-most position. To be included in the count of particles reaching $m = \xi$ at step N, the last position of the particle must be (ξ, N) *and* it may never have reached or crossed $m = \xi$. The `if` condition ensures that both conditions are met. If they are, that particular walk is accepted. Once again the various right-most positions are binned and the probability is computed. Figure 18.2.4 illustrates this process using 5000 random walks and this result is compared with the probability given by Equation 18.2.11.

□

• Example 18.2.2: Wiener process

Consider the time interval $(0, t]$ and let us subdivide it into subintervals of length Δt so that there are $t/\Delta t$ subintervals. Suppose now that a particle, initially at $x = 0$, takes a step (in one space dimension) at the times $\Delta t, 2\Delta t, \ldots$ and that the size of the step is either Δx or $-\Delta x$, with a probability of $\frac{1}{2}$ that the step is to the left or right. The position of the particle $X(t)$ at time t is a random walk, which has executed $t/\Delta t$ steps. Because the position depends on the choice of Δt and Δx, $X(t)$ depends upon t, Δt and Δx.

Mathematically we can describe this random process by

$$X(t) = \sum_{n=1}^{t/\Delta t} Z_i, \tag{18.2.12}$$

where the Z_i's are independent and identically distributed with

$$P(Z_i = \Delta x) = P(Z_i = -\Delta x) = \tfrac{1}{2}, \tag{18.2.13}$$

and $n = 1, 2, \ldots$. For each Z_i,

$$E(Z_i) = 0, \qquad \text{and} \qquad \text{Var}(Z_i) = E(Z_i^2) = (\Delta x)^2. \tag{18.2.14}$$

From Equation 18.2.12, we see that

$$E[X(t)] = 0, \qquad \text{and} \qquad \text{Var}[X(t)] = E(Z_i^2) = \frac{t}{\Delta t}\text{Var}(Z_i) = \frac{t(\Delta x)^2}{\Delta t}. \tag{18.2.15}$$

Presently we have said nothing about the relationship between Δt and Δx except that both are small. However, we cannot have just any relationship between them because the variance would be either zero or infinite. The only reasonable choice is $\Delta x = \sqrt{\Delta t}$, which makes $\text{Var}[X(t)] = t$ for all values of Δt. In the limit $\Delta t \to 0$ the random variable $X(t)$ converges into a random variable, hereafter denoted by $B(t)$, with the properties that $E[B(t)] = 0$ and $\text{Var}[B(t)] = t$. The collection of random variables $\{B(t), t > 0\}$ is a continuous process in time and called a *Wiener process*. □

Our previous example shows that Brownian motion and the Wiener process are very closely linked. Because Brownian motion occurs in so many physical and biological processes we shall focus on that motion (and the corresponding Wiener process) exclusively from now on. We define the *standard Brownian motion* (or *Wiener process*) $B(t)$ as a stochastic process that has the following properties:

1. It starts at zero: $B(0) = 0$.
2. Noting that $B(t) - B(s) \sim N(0, t-s)$, $E\{[B(t)-B(s)]^2\} = t - s$ and $\text{Var}\{[B(t) - B(s)]^2\} = 2(t-s)^2$. Replacing t with $t + dt$ and s with t, we find that $E\{[dB(t)]^2\} = dt$.
3. It has stationary and independent increments. Stationary increments means that $B(t+h) - B(\eta + h) = B(t) - B(\eta)$ for all h. An independent increment means $B(t_2) - B(t_1), \ldots, B(t_n) - B(t_{n-1})$ are independent random variables.
4. Because increments of Brownian motion on adjacent intervals are independent regardless of the *length of the interval*, the derivative will oscillate wildly as $\Delta x \to 0$ and never converge. Consequently, Brownian motion is *nowhere differentiable*.
5. It has continuous sample paths, i.e., "no jumps."
6. The expectation values for the moments are given by

$$E[B^{2n}(t)] = \frac{(2n)!t^n}{n!2^n}, \qquad \text{and} \qquad E[B^{2n-1}(t)] = 0, \tag{18.2.16}$$

where $n > 0$. See Problem 1 at the end of Section 18.4.

Problems

1. Show that $E\{\sin[aB(t)]\} = 0$, where a is a real.

2. Show that

$$E\{\cos[aB(t)]\} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} (a^2 t)^n,$$

where a is a real.

3. Show that $E\{\exp[aB(t)]\} = \exp(a^2t/2)$, where a is a real.

Project: Probabilistic Solutions of Laplace's Equation

In Section 9.7 and Section 9.8 we showed that we could solve Laplace's equation using finite difference or finite element methods, respectively. During the 1940s, the apparently unrelated fields of random processes and potential theory were shown to be in some sense mathematically equivalent.[9] As a result, it is possible to use Brownian motion to solve Laplace's equation, as you will discover in this project. The present numerical method is useful for the following reasons: (1) the entire region need not be solved in order to determine potentials at relatively few points, (2) computation time is not lengthened by complex geometries, and (3) a probabilistic potential theory computation is more topologically efficient than matrix manipulations for problems in two and three spatial dimensions.

To understand this technique,[10] consider the following potential problem:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \qquad 0 < x < 1, \quad 0 < y < 1, \tag{1}$$

subject to the boundary conditions

$$u(x, 0) = 0, \qquad u(x, 1) = x, \qquad 0 < x < 1, \tag{2}$$

and

$$u(0, y) = u(1, y) = 0, \qquad 0 < y < 1. \tag{3}$$

If we introduce a uniform grid with $\Delta x = \Delta y = \Delta s$, then the finite difference method yields the difference equation:

$$4u(i, j) = u(i + 1, j) + u(i - i, j) + u(i, j + 1) + u(i, j - 1), \tag{4}$$

with $i, j = 1, N - 1$ and $\Delta s = 1/N$.

Consider now a random process of the Markov type in which a large number N_1 of non-interacting particles are released at some point (x_1, y_1) and subsequently perform Brownian motion in steps of length Δs each unit of time. At some later time, when a few arrive at point (x, y), we define a probability $P(i, j)$ of any of them reaching the boundary $y = 1$ with potential u_k at any subsequent time in the future. Whenever one of these particles does (later) arrive on $y = 1$, it is counted and removed from the system. Because $P(i, j)$ is defined over an infinite time interval of the diffusion process, the probability of any particles leaving (x, y) and arriving along some other boundary (where the potential equals 0) at some future time is $1 - P(i, j)$. Whenever a particle arrives along these boundaries it is also removed from the square.

Having defined $P(i, j)$ for an arbitrary (x, y), we now compute it in terms of the probabilities of the neighboring points. Because the process is Markovian, where a particle jumps from a point to a neighbor with no memory of the past,

$$\begin{aligned} P(i, j) = {} & p(i + 1, j|i, j)P(i + 1, j) + p(i - 1, j|i, j)P(i - 1, j) \\ & + p(i, j + 1|i, j)P(i, j + 1) + p(i, j - 1|i, j)P(i, j - 1), \end{aligned} \tag{5}$$

---

[9] See Hersh, R., and R. J. Griego, 1969: Brownian motion and potential theory. *Sci. Amer.*, **220**, 67–74.

[10] For the general case, see Bevensee, R. M., 1973: Probabilistic potential theory applied to electrical engineering problems. *Proc. IEEE*, **61**, 423–437.

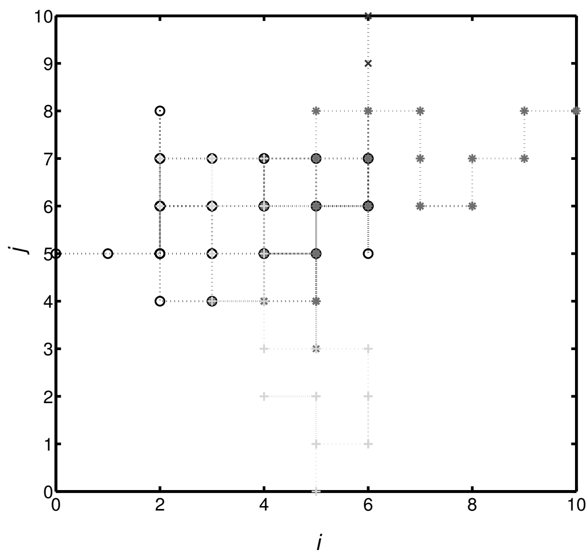


Figure 18.2.5: Four Brownian motions within a square domain with $\Delta x = \Delta y$. All of the random walks begin at grid point $i = 4$ and $j = 6$.

where $p(i+1, j|i, j)$ is the conditional probability of jumping to $(x + \Delta s, y)$, given that the particle is at (x, y). Equation (5) evaluates $P(i, j)$ as the sum of the probabilities of reaching $y = 1$ at some future time by various routes through the four neighboring points around (x, y). The sum of all the p's is exactly one because a particle at (x, y) must jump to a neighboring point during the next time interval.

Let us now compare Equation (4) and Equation (5). The potential $u(i, j)$ in Equation (4) and $P(i, j)$ becomes an identity if we take the conditional probabilities as

$$p(i+1, j|i, j) = p(i-1, j|i, j) = p(i, j+1|i, j) = p(i, j-1|i, j) = \tfrac{1}{4}, \tag{6}$$

and if we also force $u(i, N) = P(i, N) = i$, $u(i, 0) = P(i, 0) = 0$, $u(0, j) = P(0, j) = 0$, and $u(N, j) = P(N, j) = 0$. Both the potential u and the probability P become continuous functions in the space as $\Delta s \to 0$, and both are well behaved as (x, y) approaches a boundary point. A particle starting along $y = 1$, where the potential is u_k, has a probability u_k of arriving there; a particle starting on the remaining boundaries, where the potential is zero, is immediately removed with no chance of arriving along $y = 1$. From these considerations, we have

$$u(i, j) \equiv P(i, j) = \lim_{N \to \infty} \frac{1}{N} \sum_k N_k u_k, \tag{7}$$

where N is the number of particles starting at (x, y) and N_k equals the number of particles that eventually - after infinite time - arrive along the entire boundary at potential u_k. This sum includes the boundary $y = 1$ and (trivially) the remaining boundaries.

Step 1: Develop a MATLAB code to perform two-dimensional Brownian motion. Let U be a uniformly distributed random variable lying between 0 and 1. You can use `rand`. If $0 < U \le \frac{1}{4}$, take one step to the right; $\frac{1}{4} < U \le \frac{1}{2}$, take one step to the left; if $\frac{1}{2} < U \le \frac{3}{4}$, take one step downward; and if $\frac{3}{4} < U \le 1$, take one step upward. For the arbitrary point `i,j` located on a grid of $N \times N$ points with $2 \le i, j \le N-1$, repeatedly take a random step

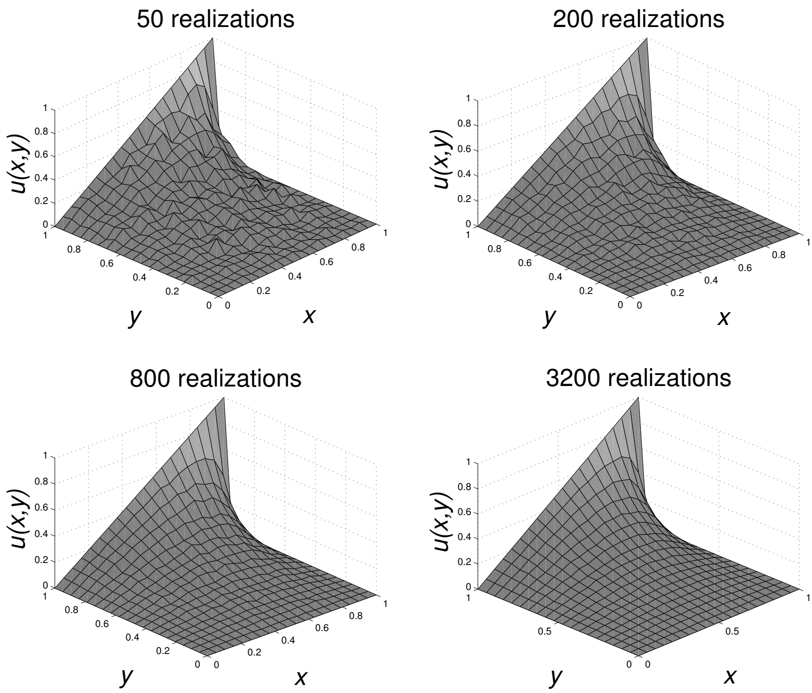


Figure 18.2.6: Solution to Equation (1) through Equation (3) using the probabilistic solution method.

until you reach one of the boundaries. Record the value of the potential at the boundary point. Let us call this result `u_k(1)`. Figure 18.2.5 illustrates four of these two-dimensional Brownian motions.

Step 2: Once you have confidence in your two-dimensional Brownian motion code, generalize it to solve Equation (1) through Equation (3) using `runs` realizations at some interior grid point. Then the solution `u(i,j)` is given by

$$\mathtt{u(i,j)} = \frac{1}{\mathtt{runs}} \sum_{n=1}^{\mathtt{runs}} \mathtt{u\_k}(n). \tag{8}$$

Step 3: Finally, plot your results. Figure 18.2.6 illustrates the potential field for different values of `runs`. What are the possible sources of error in using this method?

18.3: ITÔ'S STOCHASTIC INTEGRAL

In the previous section we noted that Brownian motion (the Wiener process) is nowhere differentiable. An obvious question is what is meant by the integral of a stochastic variable.

Consider the interval $[a, b]$, which we subdivide so that $a = t_0 < t_1 < t_2 < \cdots < t_n = b$. The earliest and simplest definition of the integral is

$$\int_a^b f(t)\,dt = \lim_{\Delta t \to 0} \sum_{i=1}^{n} f(\tau_i)\,\Delta t_i, \tag{\textbf{18.3.1}}$$

where $t_{i-1} \le \tau_i \le t_i$ and $\Delta t_i = t_i - t_{i-1}$. In the case of the classic integral, the integration is with regards to the increment dt.

Itô's integral is an integral where the infinitesimal increment involves Brownian motion $dB(t)$, which is a random variable. Before we can define this integral, we must introduce two important concepts. The first one is nonanticipating processes: A process $F(t)$ is a *nonanticipating process* if $F(t)$ is independent of any future increment $B(s) - B(t)$ for any s and t where $s > t$. Nonanticipating processes are important because Itô's integral applies only to them.

The second important concept is convergence in the mean square sense. It is defined by

$$\lim_{n\to\infty} E\left\{\left[S_n - \int_a^b F(t)\, dB(t)\right]^2\right\} = 0, \qquad \mathbf{(18.3.2)}$$

where S_n is the partial sum

$$S_n = \sum_{i=1}^{n} F(t_{i-1})\left[B(t_i) - B(t_{i-1})\right]. \qquad \mathbf{(18.3.3)}$$

We are now ready to define the Itô integral: It is the limit of the partial sum S_n:

$$\underset{n\to\infty}{\text{ms}-\lim}\, S_n = \int_a^b F(t)\, dB(t), \qquad \mathbf{(18.3.4)}$$

where we denoted the limit in the mean square sense by ms$-$lim. Combining Equation 18.3.3 and Equation 18.3.4 together, we find that

$$\int_a^b f[t, B(t)]\, dB(t) = \lim_{\Delta t\to 0} \sum_{i=1}^{n} f[t_{i-1}, B(t_{i-1})]\left[B(t_i) - B(t_{i-1})\right], \qquad \mathbf{(18.3.5)}$$

where $t_i = i\Delta t$ and $\Delta t = (b-a)/N$. As one might suspect,

$$\int_a^b dB(t) = B(b) - B(a). \qquad \mathbf{(18.3.6)}$$

Because $F(t)$ and $dB(t)$ are random variables, so is Itô's integral.

The results from Equation 18.3.6 would be misunderstood if we think about them as we do in conventional calculus. We cannot evaluate the right side of Equation 18.3.6 by looking up $B(t)$ in some book entitled "Tables of Brownian Motion." This equation only holds true for a particular realization (sample path).

• Example 18.3.1

Let us use the definition of the Itô integral to evaluate Itô integral $\int_0^t B(x)\, dB(x)$. In the present case,

$$S_n = \sum_{i=1}^{n} B(x_{i-1})\left[B(x_i) - B(x_{i-1})\right], \qquad \mathbf{(18.3.7)}$$

where $x_i = it/n$. Because $2a(b-a) = b^2 - a^2 - (b-a)^2$,

$$S_n = \frac{1}{2}\sum_{i=1}^{n} B^2(x_i) - \frac{1}{2}\sum_{i=1}^{n} B^2(x_{i-1}) - \frac{1}{2}\sum_{i=1}^{n}[B(x_i) - B(x_{i-1})]^2 \tag{18.3.8}$$

$$= \tfrac{1}{2}B^2(t) - \frac{1}{2}\sum_{i=1}^{n}[B(x_i) - B(x_{i-1})]^2. \tag{18.3.9}$$

Therefore,

$$\underset{n\to\infty}{\text{ms}-\lim} S_n = \tfrac{1}{2}B^2(t) - \tfrac{1}{2}\ \underset{n\to\infty}{\text{ms}-\lim}\sum_{i=1}^{n}[B(x_i) - B(x_{i-1})]^2 \tag{18.3.10}$$

$$= \tfrac{1}{2}B^2(t) - \frac{t}{2}. \tag{18.3.11}$$

As a consequence,

$$\int_0^t B(\eta)\,dB(\eta) = \tfrac{1}{2}B^2(t) - \frac{t}{2}, \tag{18.3.12}$$

or

$$\int_a^b B(t)\,dB(t) = \tfrac{1}{2}[B^2(b) - B^2(a)] - \frac{b-a}{2}. \tag{18.3.13}$$

Consider now the derivative of $B^2(t)$,

$$d[B^2(t)] = [B(t+dt) - B(t)]^2 = 2B(t)\,dB(t) + dB(t)\,dB(t). \tag{18.3.14}$$

In order for Equation 18.3.12 and Equation 18.3.14 to be consistent, we arrive at the very important result that

$$[dB(t)]^2 = dt \tag{18.3.15}$$

in the mean square sense. We will repeatedly use this result in the remaining portions of the chapter. □

Because the Itô integral is a random variable, two important quantities are its mean and variance. Let us turn first to the computation of the expectation of $\int_a^b f[t, B(t)]\,dB(t)$. From Equation 18.3.5 we find that

$$E\left\{\int_a^b f[t, B(t)]\,dB(t)\right\} = \lim_{\Delta t\to 0} E\left\{\sum_{i=1}^{n} f[t_{i-1}, B(t_{i-1})]\Delta B_i\right\} \tag{18.3.16}$$

$$= \lim_{\Delta t\to 0}\sum_{i=1}^{n} E\{f[t_{i-1}, B(t_{i-1})]\}E[\Delta B_i] = 0. \tag{18.3.17}$$

Therefore

$$E\left\{\int_a^b f[t, B(t)]\,dB(t)\right\} = 0. \tag{18.3.18}$$

To compute the variance, we begin by noting that

$$\left\{\int_a^b f[t,B(t)]\,dB(t)\right\}^2 = \lim_{\Delta t\to 0}\left\{\sum_{i=1}^n f[t_{i-1},B(t_{i-1})]\Delta B_i\right\}^2 \tag{18.3.19}$$

$$= \lim_{\Delta t\to 0}\sum_{i=1}^n f^2[t_{i-1},B(t_{i-1})](\Delta B_i)^2 \tag{18.3.20}$$

$$+\,2\sum_{\substack{i=1\\ i\neq j}}^n\sum_{j=1}^n f[t_{i-1},B(t_{i-1})]\Delta B_i\, f[t_{j-1},B(t_{j-1})]\Delta B_j.$$

Taking the expectation of both sides of Equation 18.3.20, we have that

$$E\left[\left\{\int_a^b f[t,B(t)]\,dB(t)\right\}^2\right] = \lim_{\Delta t\to 0}\sum_{i=1}^n E\{f^2[t_{i-1},B(t_{i-1})]\}E[(\Delta B_i)^2] \tag{18.3.21}$$

$$+\,2\lim_{\Delta t\to 0}\sum_{\substack{i=1\\ i\neq j}}^n\sum_{j=1}^n E\{f(t_{i-1},B(t_{i-1}))\}E[\Delta B_i]\,E\{f(t_{j-1},B(t_{j-1}))\}E[\Delta B_j]$$

$$= \lim_{\Delta t\to 0}\sum_{i=1}^n E\{f^2[t_{i-1},B(t_{i-1})]\}(t_{i+1}-t_i). \tag{18.3.22}$$

The double summation vanishes because of the independence of Brownian motion. Therefore, the final result is

$$E\left\{\int_a^b f[t,B(t)]\,dB(t)\right\}^2 = \int_a^b E\{f^2[t,B(t)]\}\,dt. \tag{18.3.23}$$

□

• Example 18.3.2

Consider the random number $X=\int_a^b \sqrt{t}\sin[B(t)]\,dB(t)$. Let us find $E(X)$ and $E(X^2)$.

From Equation 18.3.18, we have that $E(X)=0$. For that reason, $\mathrm{var}(X)=E(X^2)$ and

$$\mathrm{Var}(X)=E(X^2)=\int_a^b E\left\{\left|\sqrt{t}\,\sin[B(t)]\right|^2\right\}\,dt = \int_a^b t\,E\left\{\sin^2[B(t)]\right\}\,dt \tag{18.3.24}$$

$$= \int_a^b (t/2)E\left\{1-\cos[2B(t)]\right\}\,dt = \int_a^b \frac{t}{2}\left[1-\sum_{n=0}^\infty \frac{(-1)^n 2^{2n}t^n}{2^n n!}\right]\,dt \tag{18.3.25}$$

$$= -\frac{1}{2}\int_a^b \sum_{n=1}^\infty \frac{(-1)^n 2^n}{n!}t^{n+1}\,dt = \frac{1}{2}\sum_{n=1}^\infty \frac{(-1)^{n+1}2^n}{(n+2)n!}\left(b^{n+2}-a^{n+2}\right). \tag{18.3.26}$$

The value of $E\{\cos[2B(t)]\}$ follows from Problem 2 at the end of the last section. □

Table 18.3.1 gives a list of Itô stochastic integrals. Most of these results were not derived from the definition of the Itô stochastic integral but from Itô lemma, to which we now turn.

Problems

Consider the random variable $X = \int_a^b f[t, B(t)]\, dB$. Find $E(X)$ and $\text{Var}(X)$ for the following $f[t, B(t)]$:

1. $f[t, B(t)] = t$
2. $f[t, B(t)] = tB(t)$
3. $f[t, B(t)] = |B(t)|$
4. $f[f, B(t)] = \sqrt{t}\, \exp[B(t)]$
5. If $X = \int_a^b f(t)\{\sin[B(t)] + \cos[B(t)]\}\, dB(t)$, show that $\text{var}(X) = \int_a^b f^2(t)\, dt$, if $f(t)$ is a real function.

Project: Numerical Integration of Itô's Integral

Equation 18.3.5 is useful for numerically integrating the Itô integral

$$\int_0^t f[x, B(x)]\, dB(x).$$

Write a MATLAB script to check Example 18.3.1 for various values of n when $t = 1$. How does the error vary with n?

Project: Numerical Check of Equations 18.3.18 and 18.3.23

Using the script from the previous project, develop MATLAB code to compute Equation 18.3.18 and Equation 18.3.23. Using a million realizations (sample paths), compare your numerical results with the exact answer when $a = 1$, $b = 1$, and $f[t, B(t)] = \sqrt{t}\, \sin[B(t)]$.

18.4: ITÔ'S LEMMA

Before we can solve stochastic differential equations, we must derive a key result in stochastic calculus: *Itô's formula or lemma*. This is stochastic calculus's version of the chain rule.

Consider a function $f(t)$ that is twice differentiable. Using Taylor's expansion,

$$df(B) = f(B + dB) - f(B) = f'(B)\, dB + \tfrac{1}{2} f''(B)\, (dB)^2 + \cdots, \qquad \textbf{(18.4.1)}$$

where $B(t)$ denotes Brownian motion. Integrating Equation 18.4.1 from s to t, we find that

$$\int_s^t df(B) = f[B(t)] - f[B(s)] = \int_s^t f'(B)\, dB + \tfrac{1}{2} \int_s^t f''(B)\, dx + \cdots, \qquad \textbf{(18.4.2)}$$

because $[dB(x)]^2 = dx$. The first integral on the right side of Equation 18.4.2 is an Itô's stochastic integral while the second one can be interpreted as the Riemann integral of $f''(B)$. Therefore, Itô's lemma or formula is

$$f[B(t)] - f[B(s)] = \int_s^t f'(B)\, dB + \tfrac{1}{2} \int_s^t f''(B)\, dx \qquad \textbf{(18.4.3)}$$

Table 18.3.1: A Table of Itô Stochastic Integrals with $t > 0$ and $b > a > 0$

1.
$$\int_a^b dB(t) = B(b) - B(a)$$

2.
$$\int_0^t B(\eta)\, dB(\eta) = \tfrac{1}{2}[B^2(t) - t]$$

3.
$$\int_0^t [B^2(\eta) - \eta]\, dB(\eta) = \tfrac{1}{3}B^2(t) - tB(t)$$

4.
$$\int_0^t \eta\, dB(\eta) = tB(t) - \int_0^t B(\eta)\, d\eta$$

5.
$$\int_0^t B^2(\eta)\, dB(\eta) = \tfrac{1}{3}B^3(t) - \int_0^t B(\eta)\, d\eta$$

6.
$$\int_0^t e^{\lambda^2\eta/2}\cos[\lambda B(\eta)]\, dB(\eta) = \frac{1}{\lambda}e^{\lambda^2 t/2}\sin[\lambda B(t)]$$

7.
$$\int_0^t e^{\lambda^2\eta/2}\sin[\lambda B(\eta)]\, dB(\eta) = \frac{1}{\lambda}\left\{1 - e^{\lambda^2 t/2}\cos[\lambda B(t)]\right\}$$

8.
$$\int_0^t \exp\left[-\tfrac{1}{2}\lambda^2\eta \pm \lambda B(\eta)\right]\, dB(\eta) = \pm\frac{1}{\lambda}\left\{\exp\left[-\tfrac{1}{2}\lambda^2 t \pm \lambda B(t)\right] - 1\right\}$$

9.
$$\int_a^b B(\eta)\exp\left[\frac{B^2(\eta)}{2\eta}\right]\frac{dB(\eta)}{\eta^{3/2}} = b^{-1/2}\exp\left[\frac{B^2(b)}{2b}\right] - a^{-1/2}\exp\left[\frac{B^2(a)}{2a}\right]$$

10.
$$\int_a^b f(\eta)\, dB(\eta) = f(t)B(t)\Big|_a^b - \int_a^b f'(\eta)B(\eta)\, d\eta$$

11.
$$\int_a^b g'[B(\eta)]\, dB(\eta) = g[B(t)]\Big|_a^b - \frac{1}{2}\int_a^b g''[B(\eta)]\, d\eta$$

---

for $t > s$.

• Example 18.4.1

Consider the case when $f(t) = t^2$ and $s = 0$. Then, Itô's formula yields

$$B^2(t) - B^2(0) = 2\int_0^t B(x)\,dB(x) - \int_0^t dx. \tag{18.4.4}$$

Evaluating the second integral and noting that $B(0) = 0$, we again obtain Equation 18.3.12, that

$$\int_0^t B(x)\,dB(x) = \tfrac{1}{2}[B^2(t) - t]. \tag{18.4.5}$$

□

• Example 18.4.2

Consider the case when $f(t) = e^{at}$ and $s = 0$. Then, Itô's formula yields

$$e^{aB(t)} - 1 = a\int_0^t e^{aB(x)}\,dB(x) + \frac{a^2}{2}\int_0^t e^{aB(x)}\,dx. \tag{18.4.6}$$

Computing the expectation of both sides,

$$E\left[e^{aB(t)}\right] - 1 = \frac{a^2}{2}\int_0^t E\left[e^{aB(x)}\right]\,dx. \tag{18.4.7}$$

Solving this integral equation, we find that $E[e^{aB(t)}] = e^{a^2t/2}$, a result that we found earlier in Problem 2, Section 18.2. □

• Example 18.4.3

If $f(t) = \sin(\lambda t)$, $\lambda > 0$, then Itô's formula gives

$$\sin[\lambda B(t)] = \lambda\int_0^t \cos[\lambda B(\eta)]\,dB(\eta) - \tfrac{1}{2}\lambda^2\int_0^t \sin[\lambda B(\eta)]\,d\eta. \tag{18.4.8}$$

Taking the expectation of both sides of Equation 18.4.8, we find that

$$E\{\sin[\lambda B(t)]\} = -\tfrac{1}{2}\lambda^2\int_0^t E\{\sin[\lambda B(\eta)]\}\,d\eta. \tag{18.4.9}$$

Setting $g(t) = E\{\sin[\lambda B(t)]\}$, then

$$g(t) = -\tfrac{1}{2}\lambda^2\int_0^t g(\eta)\,d\eta. \tag{18.4.10}$$

The solution to this integral equation is $g(t) = 0$. Therefore, $E\{\sin[\lambda B(t)]\} = 0$. □

Educated at the Imperial University of Tokyo, Kiyoshi Itô (1915–2008) applied the techniques of differential and integral to stochastic processes. Much of Itô's original work from 1938 to 1945 was done while he worked for the Japanese National Statistical Bureau. After receiving his doctorate, Itô became a professor at the University of Kyoto from 1952 to 1979. (Author: Konrad Jacobs, Source: Archives of the Mathematisches Forschungsinstitut Oberwolfach.)

The second version of Itô's lemma begins with the second-order Taylor expansion of the function $f(t, x)$:

$$\begin{aligned} f[t+dt, B(t+dt)] - f[t, B(t)] &= f_t[t, B(t)]\,dt + f_x[t, B(t)]\,dB(t) \\ &\quad + \tfrac{1}{2}\left\{ f_{tt}[t, B(t)]\,(dt)^2 + f_{xt}[t, B(t)]\,dt\,dB(t) \right. \\ &\quad \left. + f_{xx}[t, B(t)]\,[dB(t)]^2 \right\} + \cdots . \end{aligned} \tag{18.4.11}$$

Here we assume that $f[t, B(t)]$ has continuous partial derivatives of at least second order. Neglecting higher-order terms in Equation 18.4.11, which include the terms with factors such as $(dt)^2$ and $dt\,dB(t)$ but *not* $[dB(t)]^2$ because $[dB(t)]^2 = dt$, our second version of Itô's lemma is

$$f[t, B(t)] - f[s, B(s)] = \int_s^t \left\{ f_t[\eta, B(\eta)] + \tfrac{1}{2} f_{xx}[\eta, B(\eta)] \right\} d\eta + \int_s^t f_x[\eta, B(\eta)]\,dB(\eta) \tag{18.4.12}$$

if $t > s$.

• **Example 18.4.4**

Consider the function $f(t,x) = e^{x-t/2}$. Then,

$$f_t(t,x) = -\tfrac{1}{2}e^{x-t/2}, \quad f_x(t,x) = e^{x-t/2}, \quad \text{and} \quad f_{xx}(t,x) = e^{x-t/2}. \tag{18.4.13}$$

Therefore, from Itô's lemma, we have that

$$e^{B(t)-t/2} - e^{B(s)-s/2} = \int_s^t e^{-\eta/2}e^{B(\eta)}\,dB(\eta). \tag{18.4.14}$$

□

• **Example 18.4.5: Integration by parts**

Consider the case when $F(t,x) = f(t)g(x)$. The Itô formula gives

$$d[f(t)g(x)] = \left\{f'(t)g[B(t)] + \tfrac{1}{2}f(t)g''[B(t)]\right\}\,dt + f(t)g'[B(t)]\,dB(t). \tag{18.4.15}$$

Integrating both sides of Equation 18.4.15, we find that

$$\int_a^b f(t)g'[B(t)]\,dB(t) = f(t)g[B(t)]\Big|_a^b - \int_a^b f'(t)g[B(t)]\,dt - \tfrac{1}{2}\int_a^b f(t)g''[B(t)]\,dt, \tag{18.4.16}$$

which is the stochastic version of integration by parts.

For example, let us choose $f(t) = e^{\alpha t}$ and $g(x) = \sin(x)$. Equation 18.4.16 yields

$$\int_0^t e^{\alpha\eta}\cos[B(\eta)]\,dB(\eta) = e^{\alpha\eta}\sin[B(\eta)]\Big|_0^t - \alpha\int_0^t e^{\alpha\eta}\sin[B(\eta)]\,d\eta + \tfrac{1}{2}\int_0^t e^{\alpha\eta}\sin[B(\eta)]\,d\eta \tag{18.4.17}$$

$$= e^{\alpha t}\sin[B(t)] - \left(\alpha - \tfrac{1}{2}\right)\int_0^t e^{\alpha\eta}\sin[B(\eta)]\,d\eta. \tag{18.4.18}$$

In the special case of $\alpha = \frac{1}{2}$, Equation 18.4.18 simplifies to

$$\int_0^t e^{\alpha\eta}\cos[B(\eta)]\,dB(\eta) = e^{t/2}\sin[B(t)]. \tag{18.4.19}$$

□

An important extension of Itô's lemma involves the function $f[t, X(t)]$ where $X(t)$ is no longer simply Brownian motion but is given by the first-order stochastic differential equation

$$dX(t) = cX(t)\,dt + \sigma X(t)\,dB(t), \tag{18.4.20}$$

where c and σ are real. The second-order Taylor expansion of the function $f[t, X(t)]$ becomes

$$\begin{aligned} f[t+dt, &X(t+dt)] - f[t, X(t)] = f_t[t, X(t)]\,dt + f_x[t, X(t)]\,dX(t) \\ &+ \tfrac{1}{2}\left\{f_{tt}[t, X(t)]\,(dt)^2 + f_{xt}[t, X(t)]\,dt\,dX(t) + f_{xx}[t, X(t)]\,[dX(t)]^2\right\} + \cdots. \end{aligned} \tag{18.4.21}$$

Next, we substitute for $dX(t)$ using Equation 18.4.20, neglect terms involving $(dt)^2$ and $dt\,dB(t)$, and substitute $[dB(t)]^2 = dt$. Consequently,

$$df = f[t+dt, X(t+dt)] - f[t, X(t)] \tag{18.4.22}$$

$$= \sigma X(t) f_x[t, X(t)]\, dB(t) + \left\{ f_t[t, X(t)] + cX(t) f_x[t, X(t)] + \tfrac{1}{2}\sigma^2 X^2(t) f_{xx}[t, X(t)] \right\} dt. \tag{18.4.23}$$

The present extension of Itô's lemma reads

$$f[t, X(t)] - f[s, X(s)] = \int_s^t \left\{ f_t[\eta, X(\eta)] + cX(\eta) f_x[\eta, X(\eta)] + \tfrac{1}{2}\sigma^2 X^2(\eta) f_{xx}[\eta, X(\eta)] \right\} d\eta$$
$$+ \int_s^t \sigma X(\eta) f_x[\eta, X(\eta)]\, dB(\eta) \tag{18.4.24}$$
$$= \int_s^t \left\{ f_t[\eta, X(\eta)] + \tfrac{1}{2}\sigma^2 X^2(\eta) f_{xx}[\eta, X(\eta)] \right\} d\eta$$
$$+ \int_s^t f_x[\eta, X(\eta)]\, dX(\eta), \tag{18.4.25}$$

where

$$dX(\eta) = cX(\eta)\, d\eta + \sigma X(\eta)\, dB(\eta) \tag{18.4.26}$$

and $t > s$.

We can finally generalize Itô's formula to the case of several Itô processes with respect to the *same* Brownian motion. For example, let $X(t)$ and $Y(t)$ denote two Itô processes governed by

$$dX(t) = A^{(1,1)}(t)\, dt + A^{(2,1)}(t)\, dB(t), \tag{18.4.27}$$

and

$$dY(t) = A^{(1,2)}(t)\, dt + A^{(2,2)}(t)\, dB(t). \tag{18.4.28}$$

For stochastic process $f[t, X(t), Y(t)]$, the Taylor expansion is

$$\begin{aligned} df[t, X(t), Y(t)] = {} & f_t[t, X(t), Y(t)]\, dt + f_x[t, X(t), Y(t)]\, dX(t) \\ & + f_y[t, X(t), Y(t)]\, dY(t) \\ & + \tfrac{1}{2} f_{xx}[t, X(t), Y(t)] A^{(2,1)}(t) A^{(2,1)}(t)\, dt + \tfrac{1}{2} f_{xy}[t, X(t), Y(t)] A^{(2,1)}(t) A^{(2,2)}(t)\, dt \\ & + \tfrac{1}{2} f_{yx}]t, X(t), Y(t)] A^{(2,2)}(t) A^{(2,1)}(t)\, dt + \tfrac{1}{2} f_{yy}[t, X(t), Y(t)] A^{(2,2)}(t) A^{(2,2)}(t)\, dt. \end{aligned} \tag{18.4.29}$$

• Example 18.4.6: Product rule

Consider the special case $f(t, x, y) = xy$. Then $f_t = 0$, $f_x = y$, $f_y = x$, $f_{xx} = f_{yy} = 0$, and $f_{xy} = f_{yx} = 1$. In this case, Equation 18.4.29 simplifies to

$$d[X(t)Y(t)] = Y(t)\, dX(t) + X(t)\, dY(t) + A^{(2,1)}[t, X(t), Y(t)] A^{(2,2)}[t, X(t), Y(t)]\, dt. \tag{18.4.30}$$

A very important case occurs when $A^{(2,1)}[t, X(t), Y(t)] = 0$ and $X(t) = g(t)$ is purely deterministic. In this case,

$$d[g(t)Y(t)] = Y(t)\, dg(t) + g(t)\, dY(t). \tag{18.4.31}$$

This is exactly the product rule from calculus.

Problems

1. (a) Use Equation 18.4.3 and $f(t) = t^n$ to show that

$$B^n(t) = n\int_0^t B^{n-1}(x)\, dB(x) + \frac{n(n-1)}{2}\int_0^t B^{n-2}(x)\, dx.$$

(b) Show that

$$E[B^n(t)] = \frac{n(n-1)}{2}\int_0^t E[B^{n-2}(x)]\, dx.$$

(c) Because $E[B(t)] = 0$ and $E[B^2(t)] = t$, show that

$$E\left[B^{2k+1}(t)\right] = 0, \qquad \text{and} \qquad E\left[B^{2k}(t)\right] = \frac{(2k)!}{2^k k!}t^k.$$

2. Let $f(t, x) = x^3/3 - tx$ and use Itô's formula to show that

$$\int_0^t [B^2(\eta) - \eta]\, dB(\eta) = \tfrac{1}{3}B^3(t) - t\, B(t).$$

3. If $f(x)$ is any continuously differentiable function, use Equation 18.4.29 to show that

$$\int_0^t f(\eta)\, dB(\eta) = f(t)B(t) - \int_0^t f'(\eta)B(\eta)\, d\eta.$$

4. If $f(t) = e^t$, use the previous problem to show that

$$\int_0^t e^\eta\, dB(\eta) = e^t B(t) - \int_0^t e^\eta B(\eta)\, d\eta.$$

5. Let $G(x)$ denote the antiderivative of $g(x)$. Use Equation 18.4.3 to show that

$$\int_a^b g[B(t)]\, dB(t) = G[B(t)]\Big|_a^b - \tfrac{1}{2}\int_a^b g'[B(t)]\, dt.$$

6. (a) If $g(x) = xe^x$, use Problem 5 to show that

$$\int_0^t B(\eta)e^{B(\eta)}\, dB(\eta) = [B(t) - 1]e^{B(t)} + 1 - \tfrac{1}{2}\int_0^t [B(\eta) + 1]e^{B(\eta)}\, d\eta.$$

(b) Use Equation 18.3.18 to show that

$$E\Big[B(t)e^{B(t)}\Big] = E\Big[e^{B(t)}\Big] - 1 + \tfrac{1}{2}\int_0^t \Big\{E\Big[B(\eta)e^{B(\eta)}\Big] + E\Big[e^{B(\eta)}\Big]\Big\}\, d\eta$$
$$= e^{t/2} - 1 + \tfrac{1}{2}\int_0^t \Big\{e^{\eta/2} + E\Big[B(\eta)e^{B(\eta)}\Big]\Big\}\, d\eta.$$

(c) Setting $g(t) = E\big[B(t)e^{B(t)}\big]$, use Laplace transforms to show that

$$E\Big[B(t)e^{B(t)}\Big] = te^{t/2}.$$

7. (a) If $g(x) = 1/(1+x^2)$, use Problem 5 to show that

$$\int_0^t \frac{dB(\eta)}{1+B^2(\eta)} = \arctan[B(t)] + \int_0^t \frac{B(\eta)}{[1+B^2(\eta)]^2}\, d\eta.$$

(b) Use Equation 18.3.18 to show that

$$\int_0^t E\left\{\frac{B(\eta)}{[1+B^2(\eta)]^2}\right\} d\eta = -E\left\{\arctan[B(t)]\right\}.$$

(c) Because

$$-\frac{3\sqrt{3}}{16} \le \frac{x}{(1+x^2)^2} \le \frac{3\sqrt{3}}{16}, \quad \text{or} \quad -\frac{3\sqrt{3}}{16}t \le \int_0^t \frac{B(\eta)}{[1+B^2(\eta)]^2}\, d\eta \le \frac{3\sqrt{3}}{16}t,$$

show that

$$-\frac{3\sqrt{3}}{16}t \le E\left\{\arctan[B(t)]\right\} \le \frac{3\sqrt{3}}{16}t.$$

8. If $g(x) = x/(1+x^2)$, use Problem 5 to show that

$$\int_0^t \frac{B(\eta)}{1+B^2(\eta)}\, dB(\eta) = \tfrac{1}{2}\log[1+B^2(t)] - \tfrac{1}{2}\int_0^t \frac{1-B^2(\eta)}{[1+B^2(\eta)]^2}\, d\eta.$$

9. Use integration by parts with $f(t) = e^{\beta t}$ and $g(x) = -\cos(x)$ to show that

$$\int_0^t e^{\beta\eta}\sin[B(\eta)]\, dB(\eta) = 1 - e^{\beta t}\cos[B(t)] + \big(\beta - \tfrac{1}{2}\big)\int_0^t e^{\beta\eta}\cos[B(\eta)]\, d\eta.$$

Then, take $\beta = \frac{1}{2}$ and show that

$$\int_0^t e^{\eta/2}\sin[B(\eta)]\, dB(\eta) = 1 - e^{t/2}\cos[B(t)].$$

10. Redo Example 18.4.3 and show that $E\{\cos[\lambda B(t)]\} = e^{-\lambda^2 t/2}$, $\lambda > 0$.

11. Use trigonometric double angle formulas to show that

$$\text{(a)} \qquad E\{\sin[t + \lambda B(t)]\} = e^{-\lambda^2 t/2}\sin(t),$$

and

$$\text{(b)} \qquad E\{\cos[t + \lambda B(t)]\} = e^{-\lambda^2 t/2}\cos(t),$$

when $\lambda > 0$.

12. Following Example 18.4.4 with $f(t,x) = \pm\lambda\exp\bigl(\pm\lambda x - \lambda^2 t/2\bigr)$, $\lambda > 0$, show that

$$\int_0^t \exp\left[\pm\lambda B(\eta) - \frac{\lambda^2\eta}{2}\right] dB(\eta) = \pm\frac{1}{\lambda}\left\{\exp\left[\pm\lambda B(t) - \frac{\lambda^2 t}{2}\right] - 1\right\}.$$

13. Following Example 18.4.4 with $f(t,x) = \exp(\lambda^2 t/2)\sin(\lambda x)$, $\lambda > 0$, show that

$$\int_0^t \exp\left(\frac{\lambda^2\eta}{2}\right)\cos[\lambda B(\eta)]\,dB(\eta) = \frac{1}{\lambda}\exp\left(\frac{\lambda^2 t}{2}\right)\sin[\lambda B(t)].$$

14. Following Example 18.4.4 with $f(t,x) = -\exp(\lambda^2 t/2)\cos(\lambda x)$, $\lambda > 0$, show that

$$\int_0^t \exp\left(\frac{\lambda^2\eta}{2}\right)\sin[\lambda B(\eta)]\,dB(\eta) = \frac{1}{\lambda}\left\{1 - \exp\left(\frac{\lambda^2 t}{2}\right)\cos[\lambda B(t)]\right\}.$$

15. Following Example 18.4.4 with $f(t,x) = t^{-1/2}\exp[x^2/(2t)]$, show that

$$\int_a^b B(t)\exp\left[\frac{B^2(t)}{2t}\right]\frac{dB(t)}{t^{3/2}} = b^{-1/2}\exp\left[\frac{B^2(b)}{2b}\right] - a^{-1/2}\exp\left[\frac{B^2(a)}{2a}\right].$$

16. The average of geometric Brownian motion on $[0,t]$ is defined by

$$G(t) = \frac{1}{t}\int_0^t e^{B(\eta)}\,d\eta.$$

Use the product rule to find $dG(t)$. Hint: Take the time derivative of $tG(t) = \int_0^t e^{B(\eta)}\,d\eta$.

18.5: STOCHASTIC DIFFERENTIAL EQUATIONS

We have reached the point where we can examine stochastic differential equations. Of all the possible stochastic differential equations, we will focus on Langevin's equation[11] - a

[11] Langevin, P., 1908: Sur la théorie du mouvement brownien. *C. R. Acad. Sci. Paris*, **146**, 530–530. English translation: Langevin, P., 1997: On the theory of Brownian motion. *Am. J. Phys.*, **65**, 1079–1081.

model of the velocity of Brownian particles. We will employ this model in a manner similar to that played by simple harmonic motion in the study of ordinary differential equations. It illustrates many of the aspects of stochastic differential equations without being overly complicated.

• Example 18.5.1

Before we consider the general stochastic differential equation, consider the following cases where we can make clever use of the product rule. For example, let us solve

$$dX(t) = [t + B^2(t)]\,dt + 2tB(t)\,dB(t), \qquad X(0) = X_0. \tag{18.5.1}$$

In the present case, we can find the solution by noting that

$$dX(t) = B^2(t)\,dt + t[2B(t)\,dB(t) + dt] = B^2(t)\,dt + t\,d[B^2(t)] = d[tB^2(t)]. \tag{18.5.2}$$

Integrating both sides of Equation 18.5.2, we find that the solution to Equation 18.5.1 is

$$X(t) = tB^2(t) + X_0. \tag{18.5.3}$$

Similarly, let us solve the stochastic differential equation

$$dX(t) = \frac{b - X(t)}{1-t}\,dt + dB(t), \qquad 0 \le t < 1, \tag{18.5.4}$$

with $X(0) = X_0$.

We begin by writing Equation 18.5.4 as

$$\frac{d[b - X(t)]}{1-t} + \frac{b - X(t)}{(1-t)^2}\,dt = -\frac{dB(t)}{1-t}. \tag{18.5.5}$$

Running the product rule backwards,

$$d\left[\frac{b - X(t)}{1-t}\right] = -\frac{dB(t)}{1-t}. \tag{18.5.6}$$

Integrating both sides of Equation 18.5.6 from 0 to t, we find that

$$\frac{b - X(t)}{1-t} = b - X(0) - \int_0^t \frac{dB(\eta)}{1-\eta}. \tag{18.5.7}$$

Solving for $X(t)$, we obtain the final result that

$$X(t) = b - [b - X(0)](1-t) + (1-t)\int_0^t \frac{dB(\eta)}{1-\eta}. \tag{18.5.8}$$

In the present case we cannot simplify the integral in Equation 18.5.8 and must apply numerical quadrature if we wish to have numerical values. □

In the introduction we showed that the solution to Langevin's equation:

$$dX(t) = cX(t)\,dt + \sigma\,dB(t), \qquad X(0) = X_0, \tag{18.5.9}$$

is

$$X(t) = X_0 + c\int_0^t X(\eta)\,d\eta + \sigma\int_0^t dB(\eta). \tag{18.5.10}$$

An obvious difficulty in understanding this solution is the presence of $X(s)$ in the first integral on the right side of Equation 18.5.21.

Let us approach its solution by considering the function $f(t,x) = e^{-ct}x$. Then, by Itô's lemma, Equation 18.4.16,

$$f[t,X(t)] - X(0) = \int_0^t \left\{f_t[\eta,X(\eta)] + cX(\eta)f_x[\eta,X(\eta)] + \tfrac{1}{2}\sigma^2 f_{xx}[\eta,X(\eta)]\right\}\,d\eta$$
$$+\int_0^t \sigma f_x[\eta,X(\eta)]\,dB(\eta), \tag{18.5.11}$$

because $f[0,X(0)] = X(0)$. Direct substitution of $f(t,x)$ into Equation 18.5.11 yields

$$e^{-ct}X(t) - X_0 = \sigma\int_0^t e^{-c\eta}\,dB(\eta). \tag{18.5.12}$$

Finally, solving for $X(t)$, we obtain

$$X(t) = X(0)e^{ct} + \sigma e^{ct}\int_0^t e^{-c\eta}\,dB(\eta), \tag{18.5.13}$$

an explicit expression for $X(t)$. For the special case when X_0 is constant, $X(t)$ is known as an *Ornstein-Uhlenbeck process.*[12]

An alternative derivation begins by multiplying Equation 18.5.9 by the integrating factor e^{-ct} so that the equation now reads

$$e^{-ct}\,dX(t) - ce^{-ct}X(t)\,dt = \sigma e^{-ct}\,dB(t). \tag{18.5.14}$$

Running the product rule, Equation 18.4.23, backwards, we have that

$$d\big[e^{-ct}X(t)\big] = \sigma e^{-ct}\,dB(t). \tag{18.5.15}$$

Integrating both sides of Equation 18.5.15, we obtain Equation 18.5.12.

• Example 18.5.2: Exact stochastic differential equation

Consider the stochastic differential equation

$$X(t) = X(0) + c\int_0^t X(s)\,ds + \sigma\int_0^t X(s)\,dB(s), \tag{18.5.16}$$

with $c, \sigma > 0$.

If $X(t) = f[t,B(t)]$, then by Itô's lemma, Equation 18.4.9,

$$X(t) = X(0) + \int_0^t \left\{f_t[s,B(s)] + \tfrac{1}{2}f_{xx}[s,B(s)]\right\}\,ds + \int_0^t f_x[s,B(s)]\,dB(s). \tag{18.5.17}$$

---

[12] Uhlenbeck and Ornstein, op. cit.

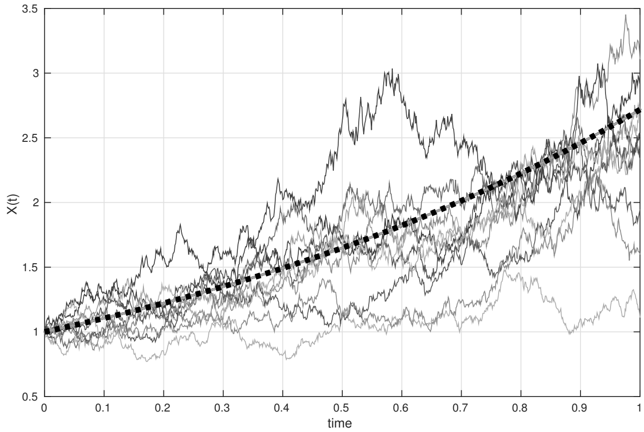


Figure 18.5.1: Ten realizations (sample paths) of geometric Brownian motion when $c = 0.1$, $\sigma = 0.5$, and $X(0) = 1$. The heavy line is the mean of $X(t)$.

Comparing Equation 18.5.16 and Equation 18.5.17, we find that

$$cf(t,x) = f_t(t,x) + \tfrac{1}{2} f_{xx}(t,x), \tag{18.5.18}$$

and

$$\sigma f(t,x) = f_x(t,x). \tag{18.5.19}$$

From Equation 18.5.19,

$$f_{xx}(t,x) = \sigma f_x(t,x) = \sigma^2 f(t,x). \tag{18.5.20}$$

Therefore, Equation 18.5.18 can be replaced by

$$\left(c - \tfrac{1}{2}\sigma^2\right) f(t,x) = f_t(t,x). \tag{18.5.21}$$

Equation 18.5.19 and Equation 18.5.21 can be solved using separation of variables, which yields

$$f(t,x) = f(0,0) \exp\left[\left(c - \tfrac{1}{2}\sigma^2\right) t + \sigma x\right], \tag{18.5.22}$$

or

$$X(t) = f[t, B(t)] = X(0) \exp\left[\left(c - \tfrac{1}{2}\sigma^2\right) t + \sigma B(t)\right]. \tag{18.5.23}$$

Thus, a stochastic differential equation can sometimes be solved as the solution of a deterministic partial differential equation. In the present case, this solution is called *geometric Brownian motion*. For its solution numerically, see Example 18.6.1. See Figure 18.5.1. □

• Example 18.5.3: Homogeneous linear equation

Consider the homogeneous linear stochastic differential equation

$$dX(t) = c_1(t)X(t)\,dt + \sigma_1(t)X(t)\,dB(t). \tag{18.5.24}$$

Let us introduce $f(t,x) = \ln(x)$. Then by Itô's lemma, Equation 18.4.21,

$$df = d[\ln(X)] = \left[c_1(t) - \tfrac{1}{2}\sigma_1^2(t)\right]\,dt + \sigma_1(t)\,dB(t), \tag{18.5.25}$$

because $f_t = 0$, $f_x = 1/x$ and $f_{xx} = -1/x^2$. Integrating both sides of Equation 18.5.25 and exponentiating the resulting expression, we obtain

$$X(t) = X(0)\exp\left[\int_0^t \left[c_1(\eta) - \tfrac{1}{2}\sigma_1^2(\eta)\right] d\eta + \int_0^t \sigma_1(\eta)\, dB(\eta)\right\}. \qquad (\mathbf{18.5.26})$$

□

• Example 18.5.4: General case

Consider the homogeneous linear stochastic differential equation

$$dX(t) = [c_1(t)X(t) + c_2(t)]\, dt + [\sigma_1(t)X(t) + \sigma_2(t)]\, dB(t). \qquad (\mathbf{18.5.27})$$

Our analysis begins by considering the homogeneous linear stochastic differential equation

$$dY(t) = c_1(t)Y(t)\, dt + \sigma_1(t)Y(t)\, dB(t), \qquad Y(0) = 1. \qquad (\mathbf{18.5.28})$$

From the previous example,

$$Y(t) = \exp\left\{\int_0^t \left[c_1(\eta) - \tfrac{1}{2}\sigma_1^2(\eta)\right] d\eta + \int_0^t \sigma_1(\eta)\, dB(\eta)\right\}. \qquad (\mathbf{18.5.29})$$

Next, let us introduce two random variables, $X_1 = 1/Y$ and $X_2 = X$. Using Itô lemma $f(t,x) = 1/x$, then

$$dX_1 = df(t,Y) = d\left(\frac{1}{Y}\right) = \left[\sigma_1^2(t) - c_1(t)\right]\frac{dt}{Y} - \sigma_1(t)\frac{dB(t)}{Y} \qquad (\mathbf{18.5.30})$$

$$= \left[\sigma_1^2(t) - c_1(t)\right] X_1(t)\, dt - \sigma_1(t)X_1(t)\, dB(t), \qquad (\mathbf{18.5.31})$$

since $f_t = 0$, $f_x = -1/x^2$ and $f_{xx} = 2/x^3$.

Using Equation 18.4.30, where X_1 is governed by Equation 18.5.31 and X_2 is governed by 18.5.27 because $X_2 = X$,

$$d(X_1X_2) = [c_2(t) - \sigma_1(t)\sigma_2(t)]\, X_1(t)\, dt + \sigma_2(t)X_1(t)\, dB(t). \qquad (\mathbf{18.5.32})$$

Upon integrating both sides of Equation 18.5.32, we have

$$X_1X_2 - X_1(0) = \int_0^t [c_2(\eta) - \sigma_1(\eta)\sigma_2(\eta)]\,\frac{d\eta}{Y(\eta)} + \int_0^t \sigma_2(\eta)\,\frac{dB(\eta)}{Y(\eta)}. \qquad (\mathbf{18.5.33})$$

Consequently, our final result is

$$X(t) = Y(t)\left\{X(0) + \int_0^t [c_2(\eta) - \sigma_1(\eta)\sigma_2(\eta)]\,\frac{d\eta}{Y(\eta)} + \int_0^t \sigma_2(\eta)\,\frac{dB(\eta)}{Y(\eta)}\right\}, \qquad (\mathbf{18.5.34})$$

where $Y(t)$ is given by Equation 18.5.29. □

- **Example 18.5.5: Stochastic Verhulst equation**

The stochastic Verhulst equation is

$$dX(t) = aX(t)[M - X(t)]\,dt + bX(t)\,dB(t), \qquad X(0) = X_0. \tag{18.5.35}$$

We begin its solution by introducing $\Phi(t) = 1/X(t)$. Then by Itô's lemma, Equation 18.4.21 with $f(x) = 1/x$,

$$d\Phi(t) = -\Phi(t)[(aM - b^2)\,dt + b\,dB(t)] + a\,dt, \qquad \Phi(0) = 1/X_0. \tag{18.5.36}$$

To solve Equation 18.5.36, we use the results from Example 18.5.4 with $c_1(t) = b^2 - aM$, $c_2(t) = a$, $\sigma_1(t) = -b$, and $\sigma_2(t) = 0$. Denoting $\epsilon(t) = (aM - b^2/2)t + bB(t)$, we can write Equation 18.5.34 as

$$\Phi(t)e^{\xi(t)} - \Phi(0) = a\int_0^t e^{\xi(\eta)}\,d\eta, \tag{18.5.37}$$

or

$$\frac{e^{\xi(t)}}{X(t)} - \frac{1}{X_0} = a\int_0^t e^{\xi(\eta)}\,d\eta. \tag{18.5.38}$$

Solving for $X(t)$, we obtain the final result that

$$X(t) = \frac{X_0\exp[\xi(t)]}{1 + aX_0\int_0^t \exp[\xi(\eta)]\,d\eta}. \tag{18.5.39}$$

Problems

1. Solve the stochastic differential equation

$$dX(t) = \tfrac{1}{2}e^{t/2}B(t)\,dt + e^{t/2}\,dB(t), \qquad X(0) = X_0,$$

by running the product rule backwards.

2. Solve the stochastic differential equation

$$dX(t) = e^{2t}[1 + 2B^2(t)]\,dt + 2e^{2t}B(t)\,dB(t), \qquad X(0) = X_0,$$

by running the product rule backwards. Hint: Rewrite the differential equation $dX(t) = e^{2t}[2B(t)\,dB(t) + dt] + (2e^{2t}\,dt)B^2(t)$.

3. Solve the stochastic differential equation

$$dX(t) = [1 + B(t)]\,dt + [t + 2B(t)]\,dB(t), \qquad X(0) = X_0,$$

by running the product rule backwards. Hint: Rewrite the differential equation $dX(t) = 2B(t)\,dB(t) + dt + B(t)\,dt + t\,dB(t)$.

4. Solve the stochastic differential equation

$$dX(t) = [3t^2 + B(t)]\,dt + t\,dB(t), \qquad X(0) = X_0,$$

by running the product rule backwards. Hint: Rewrite the differential equation $dX(t) = 3t^2\,dt + [B(t)\,dt + t\,dB(t)]$.

5. Solve the stochastic differential equation

$$dX(t) = B^2(t)\,dt + 2tB(t)\,dB(t), \qquad X(0) = X_0,$$

by running the product rule backwards. Hint: Rewrite the differential equation $dX(t) = t[2B(t)\,dB(t) + dt] + B^2(t)\,dt - t\,dt$.

6. Find the integrating factor and solution to the stochastic differential equation

$$dX(t) = [1 + 2X(t)]\,dt + e^{2t}\,dB(t), \qquad X(0) = X_0,$$

where $B(t)$ is Brownian motion.

7. Find the integrating factor and solution to the stochastic differential equation

$$dQ(t) + \frac{Q(t)}{RC}\,dt = \frac{V(t)}{R}\,dt + \frac{\alpha(t)}{R}\,dB(t), \qquad Q(0) = Q_0,$$

where R and C are real, positive constants, and $B(t)$ is Brownian motion.

8. Find the integrating factor and solution to the stochastic differential equation[13]

$$dX(t) = 2tX(t)\,dt + e^{-t}\,dt + dB(t), \qquad t \in [0, 1],$$

with $X(0) = X_0$, and $B(t)$ is Brownian motion.

9. Find the integration factor and solution to the stochastic differential equation

$$dX(t) = [4X(t) - 1]\,dt + 2\,dB(t), \qquad X(0) = X_0,$$

where $B(t)$ is Brownian motion.

10. Find the integration factor and solution to the stochastic differential equation

$$dX(t) = [2 - X(t)]\,dt + e^{-t}B(t)\,dB(t), \qquad X(0) = X_0,$$

where $B(t)$ is Brownian motion.

11. Find the integration factor and solution to the stochastic differential equation

$$dX(t) = [1 + X(t)]\,dt + e^{t}B(t)\,dB(t), \qquad X(0) = X_0,$$

where $B(t)$ is Brownian motion.

---

[13] Khodabin, M., and M. Rostami, 2015: Mean square numerical solution of stochastic differential equations by fourth order Runge-Kutta method and its applications in the electric circuits with noise. *Adv. Diff. Eq.*, **2015**, 62.

12. Find the integration factor and solution to the stochastic differential equation

$$dX(t) = \left[\tfrac{1}{2}X(t) + 1\right]\,dt + e^t\cos[B(t)]\,dB(t), \qquad X(0) = X_0,$$

where $B(t)$ is Brownian motion.

13. Find the integration factor and solution to the stochastic differential equation

$$dX(t) = \left[t + \tfrac{1}{2}X(t)\right]\,dt + e^t\sin[B(t)]\,dB(t), \qquad X(0) = X_0,$$

where $B(t)$ is Brownian motion.

14. Following Example 18.5.2, solve the exact stochastic differential equation:

$$dX(t) = e^t\left[1 + B^2(t)\right]\,dt + \left[1 + 2e^tB(t)\right]\,dB(t), \qquad X(0) = X_0.$$

Step 1: Show that $f_t + \frac{1}{2}f_{xx} = e^t(1 + x^2)$, and $f_x = 1 + 2e^tx$.

Step 2: Show that $f(t, x) = x + e^tx^2 + g(t)$.

Step 3: Show that $g(t) = X_0$ and $X(t) = B(t) + e^tB^2(t) + X_0$.

15. Following Example 18.5.2, solve the exact stochastic differential equation:

$$dX(t) = \left\{2tB^2(t) + 3t^2\left[1 + B(t)\right]\right\}\,dt + \left[1 + 3t^2B^2(t)\right]\,dB, \qquad X(0) = X_0.$$

Step 1: Show that $f_t + \frac{1}{2}f_{xx} = 2tx^3 + 3t^2(1 + x)$, and $f_x = 3t^2x^2 + 1$.

Step 2: Show that $f(t, x) = t^2x^3 + x + g(t)$.

Step 3: Show that $g'(t) = 3t^2$.

Step 4: Show that $X(t) = t^2[B^3(t) + t] + B(t) + X_0$.

Using Equation 18.5.26, solve the following stochastic differential equations:

16. $dX(t) = t^2X(t)\,dt + tX(t)\,dB(t), \qquad X(0) = X_0$
17. $dX(t) = \cos(t)X(t)\,dt + \sin(t)X(t)\,dB(t), \qquad X(0) = X_0$
18. $dX(t) = \ln(t + 1)X(t)\,dt + \sqrt{\ln(t + 1)}\,X(t)\,dB(t), \qquad X(0) = X_0$
19. $dX(t) = \ln(t + 1)X(t)\,dt + tX(t)\,dB(t), \qquad X(0) = X_0$

20. Following Example 18.5.5, solve the stochastic differential equation

$$dX(t) = [aX^n(t) + bX(t)]\,dt + cX(t)\,dB(t), \qquad X(0) = X_0,$$

where $n > 1$.

Step 1: Setting $\Phi(t) = X^{1-n}(t)$, use Itô's lemma Equation 18.4.21 with $f(x) = 1/x^{n-1}$ to show that

$$d\Phi(t) = (1 - n)\Phi(t)\left[\left(b - \tfrac{1}{2}nc^2\right)\,dt + c\,dB(t)\right] + (1 - n)a\,dt.$$

Step 2: Setting $c_1(t) = (1-n)b - n(1-n)c^2/2$, $c_2(t) = (1-n)a$, $\sigma_1(t) = (1-n)c$, and $\sigma_2(t) = 0$, show that

$$\frac{\exp[(n-1)\xi(t)]}{X^{n-1}(t)} - \frac{1}{X_0^{n-1}} = (1-n)a \int_0^t \exp[(n-1)\xi(\eta)]\, d\eta,$$

or

$$\frac{\exp[(n-1)\xi(t)]}{X^{n-1}(t)} = \frac{1}{X_0^{n-1}} + (1-n)a \int_0^t \exp[(n-1)\xi(\eta)]\, d\eta,$$

where $\xi(t) = (b - c^2/2)t + cB(t)$.

21. Following Example 18.5.5, solve the stochastic Ginzburg-Landau equation:

$$dX(t) = \left[ae^{cX(t)} + b\right] dt + \sigma\, dB(t), \qquad X(0) = X_0.$$

Step 1: Setting $\Phi(t) = \exp[-cX(t)]$, use Itô's lemma Equation 18.4.21 with $f(x) = e^{-cx}$ to show that

$$d\Phi(t) = -\left(bc - \tfrac{1}{2}\sigma^2 c^2\right) \Phi(t)\, dt - \sigma c \Phi(t)\, dB(t) - ac\, dt.$$

Step 2: Setting $c_1(t) = \sigma^2 c^2/2 - bc$, $c_2(t) = -ac$, $\sigma_1(t) = -\sigma c$, and $\sigma_2(t) = 0$, show that

$$X(t) = X_0 + bt + \sigma B(t) - \frac{1}{c} \ln\left\{1 - ac \int_0^t \exp\left[cX_0 + bc\xi + \sigma c B(\xi)\right]\, d\xi\right\}.$$

22. Following Example 18.5.5, solve the stochastic differential equation:

$$dX(t) = \left\{[1 + X(t)][1 + X^2(t)]\right] dt + [1 + X^2(t)]\, dB(t), \qquad X(0) = X_0.$$

Step 1: Setting $\Phi(t) = \tan^{-1}[X(t)]$, use Itô's lemma Equation 18.4.21 with $f(x) = \tan^{-1}(x)$ to show that $d\Phi(t) = dt + dB(t)$.

Step 2: Solving the stochastic differential equation in Step 1, show that

$$X(t) = \tan[\tan^{-1}(X_0) + t + B(t)].$$

18.6: NUMERICAL SOLUTION OF STOCHASTIC DIFFERENTIAL EQUATIONS

In this section we construct numerical schemes for integrating the stochastic differential equation

$$dX(t) = a[X(t), t]\, dt + b[X(t), t]\, dB(t) \qquad \textbf{(18.6.1)}$$

on $t_0 \le t \le T$ with the initial-value $X(t_0) = X_0$.

Our derivation begins by introducing the grid $t_0 < t_1 < t_2 < \cdots < t_n < \cdots < t_N = T$. For simplicity we assume that all of the time increments are the same and equal to $0 < \Delta t < 1$ although our results can be easily generalized when this is not true. Now

$$X_{n+1} = X_n + \int_{t_n}^{t_{n+1}} a[X(\eta), \eta]\, d\eta + \int_{t_n}^{t_{n+1}} b[X(\eta), \eta]\, dB(\eta). \qquad \textbf{(18.6.2)}$$

The crudest approximation to the integrals in Equation 18.6.2 is

$$\int_{t_n}^{t_{n+1}} a[X(\eta),\eta]\,d\eta \approx a[X(t_n),t_n]\Delta t_n, \tag{18.6.3}$$

and

$$\int_{t_n}^{t_{n+1}} b[X(\eta),\eta]\,dB(\eta) \approx b[X(t_n),t_n]\Delta B_n. \tag{18.6.4}$$

Substituting these approximations into Equation 18.6.2 yields the *Euler-Marugama approximation*.[14] For the Itô process $X(t) = \{X(t), t_0 \le t \le T\}$:

$$X_{n+1} = X_n + a(t_n, X_n)\,(t_{n+1} - t_n) + b(t_n, X_n)\left(B_{t_{n+1}} - B_{t_n}\right) \tag{18.6.5}$$

for $n = 0, 1, 2, \ldots, N-1$ with the initial value X_0.

When $b = 0$, the stochastic iterative scheme reduces to the conventional Euler scheme for ordinary differential equations. See Section 1.7. When $b \neq 0$, we have an extra term generated by the random increment $\Delta B_n = B(t_{n+1}) - B(t_n)$ where $n = 0, 1, 2, \ldots, N-1$ for Brownian motion (the Wiener process) $B(t) = B(t), t \ge 0$. Because these increments are independent Gaussian random variables, the mean equals $E(\Delta B_n) = 0$ while the variance is $E[(\Delta B_n)^2] = \Delta t$. We can generate ΔB_n using the MATLAB function `randn`.

An important consideration in the use of any numerical scheme is the rate of convergence. During the numerical simulation of a realization, at time t there will be a difference between the exact solution $X(t)$ and the numerical approximation $Y(t)$. This difference $e(t) = X(t) - Y(t)$ will also be a random variable. A stochastic differential equation scheme *converges strongly with order* m, if for any time t, $E(|e(t)|) = O[(\Delta t)^m]$ for sufficiently small time step Δt. The strong order for the Euler-Marugama method can be proven to be $\frac{1}{2}$.

To construct a strong order 1 approximation to Equation 18.6.1, we return to Equation 18.6.2. Using Equation 18.4.12, we have

$$\begin{aligned} X_{n+1} - X_n = &\int_{t_n}^{t_{n+1}} \left[a[X_n(\eta),\eta] + \int_{t_n}^{\eta} \left(a a_x + \tfrac{1}{2} b^2 a_{xx}\right)\,d\xi + \int_{t_n}^{\eta} b a_x\,dB(\xi)\right]\,d\eta \\ &+ \int_{t_n}^{t_{n+1}} \left[b[X_n(\eta),\eta] + \int_{t_n}^{\eta} \left(a b_x + \tfrac{1}{2} b^2 b_{xx}\right)\,d\xi + \int_{t_n}^{\eta} b b_x\,dB(\xi)\right]\,d\eta \qquad (18.6.6) \\ = &\, a[X(t_n),t_n]\Delta t + b[X(t_n),t_n)\Delta B_n + R_n, \qquad (18.6.7) \end{aligned}$$

where

$$R_n = \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{\eta} b b_x\,dB(\xi)\right]\,dB(\eta) + \text{higher} - /\text{order terms}. \tag{18.6.8}$$

Dropping the higher-order terms,

$$R_n \approx b[X(t_n),t_n] b_x[X(t_n),t_n] \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{\eta} dB(\xi)\right]\,dB(\eta). \tag{18.6.9}$$

Consider now the double integrals

$$(\Delta B_n)^2 = \left(\int_{t_n}^{t_{n+1}} dB(\eta)\right)\left(\int_{t_n}^{t_{n+1}} dB(\eta)\right) = \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{t_{n+1}} dB(\xi)\right]\,dB(\eta). \tag{18.6.10}$$

[14] Maruyama, G., 1955: Continuous Markov processes and stochastic equations. *Rend. Circ. Math. Palermo, Ser. 2,*, **4**, 48–90.

Now,

$$\int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{t_{n+1}} dB(\xi)\right] dB(\eta) = \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{\eta} dB(\xi)\right] dB(\eta) + \int_{t_n}^{t_{n+1}} \left[\int_{\eta}^{t_{n+1}} dB(\xi)\right] dB(\eta) + \int_{t_n}^{t_{n+1}} [dB(\eta)]^2 \quad \textbf{(18.6.11)}$$

$$= 2\int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{\eta} dB(\xi)\right] dB(\eta) + \int_{t_n}^{t_{n+1}} [dB(\eta)]^2 \quad \textbf{(18.6.12)}$$

$$= 2\int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{\eta} dB(\xi)\right] dB(\eta) + \Delta t, \quad \textbf{(18.6.13)}$$

because

$$\int_{t_n}^{t_{n+1}} [dB(\eta)]^2 = \int_{t_n}^{t_{n+1}} d\eta = \Delta t. \quad \textbf{(18.6.14)}$$

Combining Equation 18.6.9, Equation 18.6.10, and Equation 18.6.13 yields

$$R_n \approx b[X(t_n), t_n] b_x[X(t_n), t_n] \left[(\Delta B_n)^2 - \Delta t\right]. \quad \textbf{(18.6.15)}$$

Finally, substituting Equation 18.6.15 into Equation 18.6.7 gives the final result, the Milstein method:[15]

$$X_{n+1} = X_n + a(X_n, t_n)\,\Delta t_n + b(X_n, t_n)\,\Delta B_n + \tfrac{1}{2} b(X_n, t_n) \frac{\partial b(X_n, t_n)}{\partial x} \left[(\Delta B_n)^2 - \Delta t\right]. \quad \textbf{(18.6.16)}$$

• **Example 18.6.1**

Consider the Itô process $X(t)$ defined by the linear stochastic differential equation

$$dX(t) = aX(t)\,dt + bX(t)\,dB(t), \quad \textbf{(18.6.17)}$$

for $t \in [0, T]$. If this Itô process has the drift $a(x, t) = ax$ and the diffusion coefficient $b(x, t) = bx$, the exact solution (see Equation 18.5.16) is

$$X(t) = X_0 \exp\left[\left(a - \frac{b^2}{2}\right) t + bB(t)\right] \quad \textbf{(18.6.18)}$$

for $t \in [0, T]$. Figure 18.6.1 compares the numerical solution of this stochastic differential equation using the Euler-Marugama and Milstein method against the exact solution. Note that each frame has a different solution because the Brownian forcing changes with each realization. □

Although a plot of various realizations can give an idea of how the stochastic processes affect the solution, two more useful parameters are the sample mean and standard deviation at time t_n:

$$\overline{X}(t_n) = \frac{1}{J} \sum_{j=1}^{J} X_j(t_n), \quad \textbf{(18.6.19)}$$

[15] Milstein, G., 1974: Approximate integration of stochastic differential equations. *Theory Prob. Applic.*, **19**, 557–562.

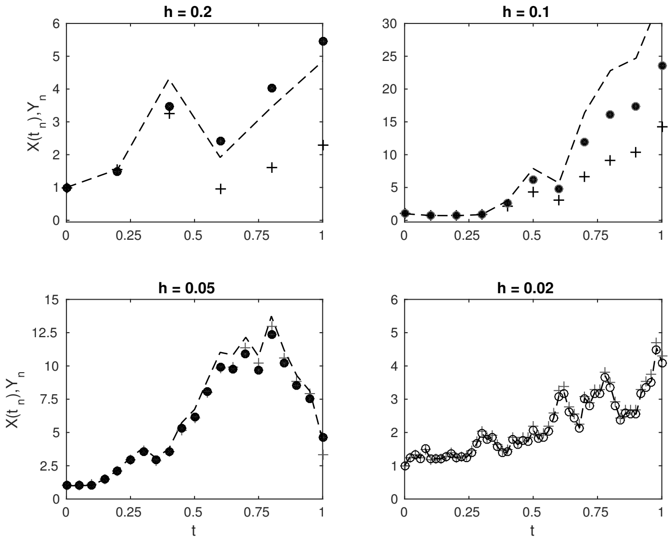


Figure 18.6.1: The numerical solution of the stochastic differential equation, Equation 18.6.17, using the Euler-Marugama (crosses) and the Milstein (circles) methods for various time steps h. The dashed line gives the exact solution.

and

$$\sigma^2(t_n) = \frac{1}{J-1} \sum_{j=1}^{J} \left[X_j(t_n) - \overline{X}(t_n) \right]^2, \tag{18.6.20}$$

where J are the number of realizations and $X_j(t_n)$ is the value of the random variable at time t_n of the jth realization.

In many physical problems, "noise" is the origin of the stochastic process and we suspect that we have a normal distribution $N(\mu, \sigma^2)$ where μ and σ are the population mean and standard deviation, respectively. Then, using the sample statistics, Equations 18.6.20 and 18.6.21, a two-sided confidence interval can be determined as

$$\left[\overline{X}(t_n) - \tau_{student} \frac{\sigma(t_n)}{\sqrt{J}}, \overline{X}(t_n) + \tau_{student} \frac{\sigma(t_n)}{\sqrt{J}} \right]$$

based on the student-τ distribution with $J-1$ degrees of freedom.

Project: RL Electrical Circuit with Noise

An important component of contemporary modelling is the mixture of deterministic and stochastic aspects of a physical system. In this project you will see how this is done using a simple electrical system.[16]

Consider a simple electrical circuit consisting of a resistor with resistance R and an inductor with inductance L. If the circuit is driven by a voltage source $v(t)$, the current $I(t)$ at a given time t is given by the first-order ordinary differential equation

$$L \frac{dI}{dt} + RI = v(t), \qquad I(0) = I_0. \tag{1}$$

[16] See Kolářová, E., 2005: Modeling RL electrical circuits by stochastic differential equations. *Proc. Int. Conf. Computers as Tool*, Belgrade (Serbia and Montenegro), **IEEE R8**, 1236–1238.

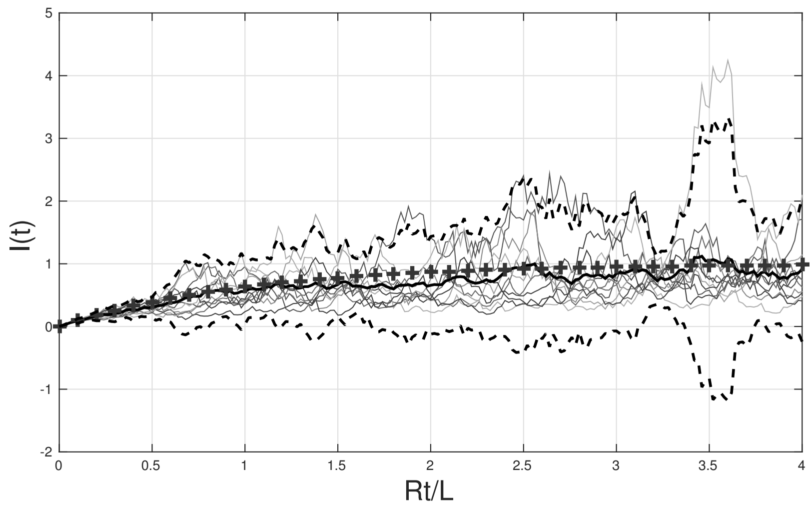


Figure 18.6.2: Eleven realizations as a function of the nondimensional time Rt/L of the numerical solution of Equation (4) using the Euler-Marugama method when $h = 0.02$, $\alpha/L = 1$, $\beta/L = 0$, $I_0 = 0$, and $v(t) = R$. The mean and 95% confidence interval (here $t_{student} = 2.228$) are given by the heavy solid and dashed lines, respectively. Finally, the crosses (+) give the deterministic solution.

Step 1: Using classical methods, show that the deterministic solution to Equation (1) is

$$I(t) = I_0 e^{-Rt/L} + \frac{1}{L}\int_0^t \exp\left[\frac{R}{L}(\tau - t)\right] v(\tau)\, d\tau. \qquad (\mathbf{2})$$

There are two possible ways that randomness can enter this problem. First, the power supply could introduce some randomness so that the right side of Equation (1) could read $v(t)+\alpha\, dB_2(t)/dt$. Second, some physical process within the resistor could cause randomness so that the resistance would now equal $R + \beta\, dB_1(t)/dt$. Here $B_1(t)$ and $B_2(t)$ denote two independent white noise processes and α, β are nonnegative constants. In this case the governing differential equation would now read

$$\frac{dI}{dt} + \frac{1}{L}\left[R + \alpha\frac{dB_1}{dt}\right] = \frac{1}{L}\left[v(t) + \beta\frac{dB_2}{dt}\right], \qquad I(0) = I_0. \qquad (\mathbf{3})$$

Converting Equation (3) into the standard form of a stochastic ordinary differential equation, we have that

$$dI = \frac{1}{L}\left[v(t) - RI(t)\right] dt - \frac{\alpha}{L} I(t)\, dB_1(t) + \frac{\beta}{L}\, dB_2(t), \qquad I(0) = I_0. \qquad (\mathbf{4})$$

Step 2: Using MATLAB, create a script to numerically integrate Equation (4) for a given set of α, β, $I_0 = 0$, R, L, and $v(t)$. Plot $I(t)$ as a function of the nondimensional time Rt/L for many realizations (say 20). See Figure 18.6.2.

Step 3: Although some idea of the effect of randomness is achieved by plotting several realizations, a better way would be to compute the mean and standard deviation at a given time. On the plot from the previous step, plot the mean and standard deviation of your

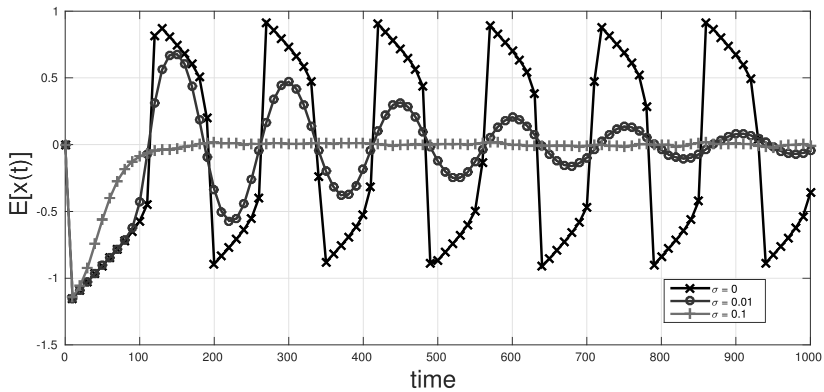


Figure 18.6.3: Plot of $E[x(t)]$ versus time for the FitzHugh-Nagamo model for three values of σ. The value of the parameters are $a = 0.8$, $m = 1.2$, and $\tau = 100$. The Euler method was used with a time step of 0.1.

solution as a function of nondimensional time. How does it compare to the deterministic solution?

Project: Relaxation Oscillator with Brownian Motion Forcing

The FitzHugh-Nagamo[17] model describes excitable systems such as a neuron. We will modify it so that the forcing is due to Brown motion. The governing equations are

$$dx = -x(x^2 - a^2)\,dt - y\,dt + \sigma\,dB_1(t),$$

and

$$dy = (x - my)\,dt/\tau + \sigma\,dB_2(t),$$

where a, m, σ, and τ are parameters.

Write a MATLAB script to numerically integrate this modified FitzHugh-Nagamo model for various values of σ. Using many simulations, compute $E[x(t)]$ as a function of time t. See Figure 18.6.3. What is the effect of the Brownian motion forcing?

Project: Stochastic Damped Harmonic Oscillator

The damped stochastic harmonic oscillator is governed by the stochastic differential equations:

$$dv(t) = -\gamma v(t)\,dt - k^2 x(t)\,dt - \alpha x(t)\,dB(t), \qquad \text{and} \qquad dx(t) = v(t)\,dt, \tag{1}$$

where k, α and γ are real constants. This system of equations is of interest for two reasons: (1) The system is forced by Brownian motion. (2) The noise is multiplicative rather than additive because the forcing term is $x(t)\,dB(t)$ rather than just $dB(t)$.

[17] FitzHugh, R., 1961: Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.*, **1**, 445-466; Nagumo, J., S. Arimoto, and S. Yoshizawa, 1962: An active pulse transmission line simulating nerve axon. *Proc. IRE*, **50**, 2061–2070.

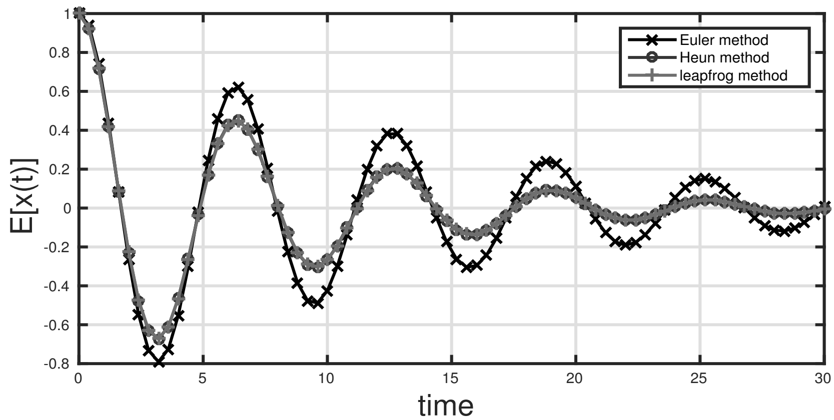


Figure 18.6.4: Plot of $E[x(t)]$ versus time for the damped harmonic oscillator forced by Brownian motion. The value of the parameters are $k = 1$, $\gamma = 0.25$, and $alpha = \Delta t = 0.1$. Five thousand realization were performed.

We could solve both equations numerically using Euler's method.[18] The purpose of this project is to introduce you to the Heun method. In the Heun method we first compute an estimate of the solution x^* and v^* by taking an Euler-like time step:

$$x^* = x_i + v_i\Delta t, \qquad \text{and} \qquad v^* = v_i - \gamma v_i \Delta t - k^2 x_i \Delta t - \alpha x_i \Delta B_i, \tag{2}$$

where x_i and v_i denote the displacement and velocity at time $t_i = i\Delta t$, Δt is the time step, and $i = 0, 1, 2, \ldots$. With these estimates we compute the value for x_{i+1} and v_{i+1} using

$$x_{i+1} = x_i + \tfrac{1}{2}(v_i + v^*)\Delta t, \qquad \text{and} \qquad v_{i+1} = v_i - \tfrac{1}{2}\gamma(v_i + v^*)\Delta t - \tfrac{1}{2}k^2(x_i + x^*)\Delta t - \alpha x_i \Delta B_i. \tag{3}$$

Qiang and Habib[19] developed a leapfrog algorithm to solve this problem. Because the algorithm is rather complicated, the interested student is referred to their paper.

Write a MATLAB script to use the Euler and Huen methods to numerically integrate the stochastic harmonic oscillator when $10\alpha = 4\gamma = k = 1$ and $x(0) = v(0) = 0$. Using many simulations, compute $E[x(t)]$ as a function of time t. See Figure 18.6.4. What happens to the accuracy of the solution for larger values of Δt?

Project: Mean First Passage Time

The stochastic differential equation

$$dX(t) = [X(t) - X^3(t)]\, dt + \frac{4}{X^2(t) + 1}\, dB(t)$$

describes the motion of a particle in a double-well potential $V(x) = x^4/4 - x^2/2$, subject to a spacially dependent random forcing when the acceleration $X''(t)$ can be neglected.

[18] For further details, see Greiner, A., W. Strittmatter, and J. Honerkamp, 1988: Numerical integration of stochastic differential equations. *J. Stat. Phys.*, **51**, 95–108.

[19] Qiang, J., and S. Habib, 2000: Second-order stochastic leapfrog algorithm for multiplicative noise Brownian motion. *Phys. Review*, **62**, 7430–7437.

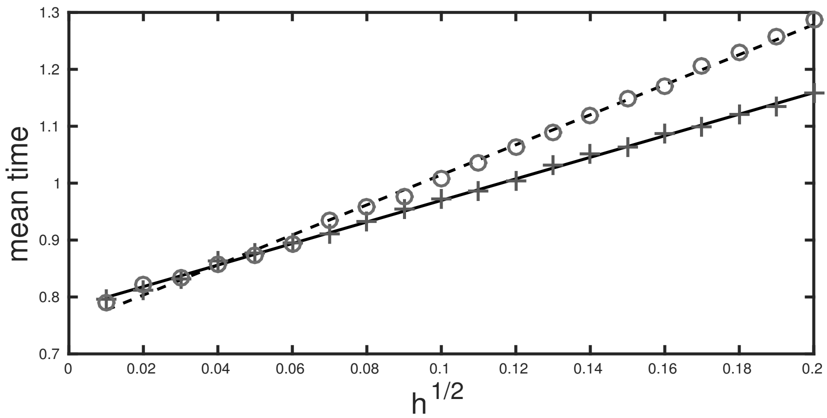


Figure 18.6.5: The mean time that it takes a particle to travel from $X(0) = -1$ to $X = 0$ in the double-well potential stated in the text. Sixty thousand realizations were used with a time step h. Two differential numerical schemes were used: the Euler-Marugama (crosses) and the Milstein (circles) methods. The curves are linear least-squares fits through the results.

An important question is what is the average (mean) time that it takes a particle initially located at a minimum $X(0) = -1$ to reach the local maximum $X(t) = 0$.

Write MATLAB code that computes $X(t)$ as a function of time t. Using this code and creating N realizations, compute the length of time that it takes the particle to reach $X(t) = 0$ in each realization. Then compute the mean from those times and plot the results as a function $\sqrt{h}$, the square root of the time step. See Figure 18.6.5. We used $\sqrt{t}$ rather than h following the suggestions of Seeßelberg and Petruccione.[20]

Project: Bankruptcy of a Company

The stochastic differential equation[21]

$$dX(t) = [\mu X(t) - iD]\,dt + \sigma X(t)\,dB(t), \qquad 0 < t < T,$$

with $X(0) = X_0$, describes the evolution with time t of the wealth $X(t)$ of a firm. Here μ and σ denote the deterministic and stochastic evolution of the firm's wealth, respectively, X_0 is the initial wealth of the firm, and iD gives the amount of payment to a financier (bank) who initially loaned the firm the amount D at the interested rate i. Write a MATLAB code to simulate the wealth of a firm during its lifetime T given a known D, i and X_0 with $\mu = 1.001$/year, and various values of σ.

During the simulation there is a chance that the firm goes bankrupt at time $t = \tau < T$. This occurs when the stochastic process hits the barrier $X(\tau) = 0$. If n denotes the number of times that bankruptcy occurs in N simulations, the probability of bankruptcy is $P[X(\tau) = 0] = n/N$. Using your code for simulating a firm's wealth, compute the probability of bankruptcy as a function of interest rate for a small ($D = 20$, $X_0 = 100$),

---

[20] Seeßelberg, M., and F. Petruccione, 1993: An improved algorithm for the estimation of the mean first passage of ordinary stochastic differential equations. *Comput. Phys. Commun.*, **74**, 247–255.

[21] See Cerqueti, R., and A. G. Quaranta, 2012: The perspective of a bank in granting credit: An optimization model. *Optim. Lett.*, **6**, 867–882.

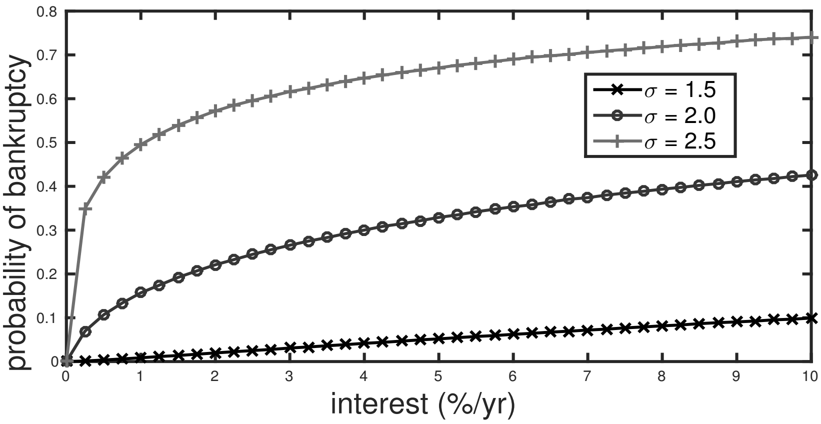


Figure 18.6.6: The probability of bankruptcy over a three-year period as a function of interest rate of a firm with initial wealth $X_0 = 500$ and debt $D = 100$. Other parameters are $h = 0.01$ yr and $\mu = 1.001$/year. The units on σ is year$^{-1/2}$. Five hundred thousand realizations were used to compute the probability.

medium ($D = 100$, $X_0 = 500$), and large ($D = 200$, $X_0 = 1000$) firm. See Figure 18.6.6. How does the average value of τ vary with interest rate?

Further Readings

Kloeden, P. E., and E. Platen, 1992: *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, 632 pp. A solid book covering numerical schemes for solving stochastic differential equations.

Mikosch, T., 1998: *Elementary Stochastic Calculus with Finance in View*. World Scientific, 212 pp. Very well-crafted book on stochastic calculus.

Answers
To the Odd-Numbered Problems

Section 1.1

1. first-order, linear
3. first-order, nonlinear
5. second-order, linear
7. third-order, nonlinear
9. second-order, nonlinear
11. first-order, nonlinear
13. first-order, nonlinear
15. second-order, nonlinear

Section 1.2

1. $y(x) = -\ln(C - x^2/2)$
3. $y^2(x) - \ln^2(x) = 2C$
5. $2 + y^2(x) = C(1 + x^2)$
7. $y(x) = -\ln(C - e^x)$
9. $\dfrac{ay^3(t) - b}{ay_0^3 - b} = e^{-3at}$
13. $V(t) = \dfrac{V_0 S\, e^{-t/(RC)}}{S + RV_0\left[1 - e^{-t/(RC)}\right]}$
15. $N(t) = N(0)\exp\bigl\{\ln[K/N(0)]\left(1 - e^{-bt}\right)\bigr\}$
17.

$$\frac{1}{([\mathrm{A}]_0 - [\mathrm{B}]_0)([\mathrm{A}]_0 - [\mathrm{C}]_0)} \ln\left(\frac{[\mathrm{A}]_0}{[\mathrm{A}]_0 - [\mathrm{X}]}\right) + \frac{1}{([\mathrm{B}]_0 - [\mathrm{A}]_0)([\mathrm{B}]_0 - [\mathrm{C}]_0)} \ln\left(\frac{[\mathrm{B}]_0}{[\mathrm{B}]_0 - [\mathrm{X}]}\right) + \frac{1}{([\mathrm{C}]_0 - [\mathrm{A}]_0)([\mathrm{C}]_0 - [\mathrm{B}]_0)} \ln\left(\frac{[\mathrm{C}]_0}{[\mathrm{C}]_0 - [\mathrm{X}]}\right) = kt$$

Section 1.3

1. $\ln|y| - x/y = C$
3. $|x|(x^2 + 3y^2) = C$
5. $y = x\left(\ln|x| + C\right)^2$
7. $\sin(y/x) - \ln|x| = C$

Section 1.4

1. $xy^2 - \frac{1}{3}x^3 = C$
3. $xy^2 - x + \cos(y) = C$
5. $y/x + \ln(y) = C$
7. $\cos(xy) = C$
9. $x^2y^3 + x^5y + y = C$
11. $xy\ln(y) + e^x - e^{-y} = C$
13. $y - x + \frac{1}{2}\sin(2x + 2y) = C$

Section 1.5

1. $y(x) = \frac{1}{2}e^x + Ce^{-x}, \quad x \in (-\infty, \infty)$
3. $y(x) = \ln(x)/x + Cx^{-1}, \quad x \neq 0$
5. $y(x) = 2x^3\ln(x) + Cx^3, \qquad x \in (-\infty, \infty)$
7. $e^{\sin(2x)}y(x) = C, \quad n\pi + \varphi < 2x < (n+1)\pi + \varphi$, where φ is any real and n is any integer.
9. $y(x) = \frac{4}{3} + \frac{11}{3}e^{-3x}, \qquad x \in (-\infty, \infty)$
11. $y(x) = (x + C)\csc(x)$
13. $y(x) = \dfrac{\cos^a(x)\ y(0)}{[\sec(x) + \tan(x)]^b} + \dfrac{c\ \cos^a(x)}{[\sec(x) + \tan(x)]^b}\int_0^x \dfrac{[\sec(\xi) + \tan(\xi)]^b}{\cos^{a+1}(\xi)}\,d\xi$
15. $y(x) = \dfrac{2ax - 1}{8a^2} + \dfrac{\omega^2 e^{-2ax}}{8a^2(a^2 + \omega^2)} - \dfrac{a\sin(2\omega x) - \omega\cos(2\omega x)}{8\omega(a^2 + \omega^2)}$
17. $y^2(x) = 2(x - x^{2/k})/(2 - k) \quad \text{if} \quad k \neq 2; \quad y^2(x) = x\ln(1/x) \quad \text{if} \quad k = 2$
19. $[\mathrm{A}](t) = [\mathrm{A}]_0 e^{-k_1 t}, \qquad [\mathrm{B}](t) = \dfrac{k_1[\mathrm{A}]_0}{k_2 - k_1}\left[e^{-k_1 t} - e^{-k_2 t}\right],$

    $[\mathrm{C}](t) = [\mathrm{A}]_0\left(1 + \dfrac{k_1 e^{-k_2 t} - k_2 e^{-k_1 t}}{k_2 - k_1}\right)$
21. $y(x) = [Cx + x\ln(x)]^{-1}$
23. $y(x) = \left[Cx^2 + \frac{1}{2}x^2\ln(x)\right]^2$
25. $y(x) = [Cx - x\ln(x)]^{1/2}$

Section 1.6

5. The equilibrium points are $x = 0, \frac{1}{2}$, and 1. The equilibrium at $x = \frac{1}{2}$ is unstable while the equilibriums at $x = 0$ and 1 are stable.

7. The equilibrium point for this differential equation is $x = 0$, which is stable.

Section 1.7

1. $x(t) = e^t + t + 1$
3. $x(t) = [1 - \ln(t + 1)]^{-1}$

Section 2.0

1. $y_2(x) = A/x$
3. $y_2(x) = Ax^{-4}$
5. $y_2(x) = A(x^2 - x + 1)$
7. $y_2(x) = A\sin(x)/\sqrt{x}$
9. $y(x) = C_2 e^{C_1 x}$ o
11. $y(x) = \left(1 + C_2 e^{C_1 x}\right)/C_1$
13. $y(x) = -\ln|1 - x|$
15. $y(x) = C_1 - 2\ln(x^2 + C_2)$

Section 2.1

1. $y(x) = C_1 e^{-x} + C_2 e^{-5x}$

3. $y(x) = C_1 e^{x} + C_2 x e^{x}$

5. $y(x) = C_1 e^{2x} \cos(2x) + C_2 e^{2x} \sin(2x)$

7. $y(x) = C_1 e^{-10x} + C_2 e^{4x}$

9. $y(x) = e^{-4x} [C_1 \cos(3x) + C_2 \sin(3x)]$

11. $y(x) = C_1 e^{-4x} + C_2 x e^{-4x}$

13. $y(x) = C_1 + C_2 x + C_3 \cos(2x) + C_4 \sin(2x)$

15. $y(x) = C_1 e^{2x} + C_2 e^{-x} \cos(\sqrt{3}\, x) + C_3 e^{-x} \sin(\sqrt{3}\, x)$

17. $y(t) = e^{-t/(2\tau)} \left\{ A \exp\left[t\sqrt{1-2A\tau}/(2\tau)\right] + B \exp\left[-t\sqrt{1-2A\tau}/(2\tau)\right] \right\}$

Section 2.2

1. $x(t) = 2\sqrt{26}\, \sin(5t + 1.7682)$

3. $x(t) = 2\cos(\pi t - \pi/3)$

5. $x(t) = s_0 \cos(\omega t) + \dfrac{v_0}{\omega} \sin(\omega t)$ and $v(t) = v_0 \cos(\omega t) - \omega s_0 \sin(\omega t)$, where $\omega^2 = Mg/(mL)$.

Section 2.3

1. $x(t) = 4e^{-2t} - 2e^{-4t}$

3. $x(t) = e^{-5t/2} \left[4\cos(6t) + \frac{13}{3} \sin(6t)\right]$

5. The roots are equal when $c = 4$ when $m = -2$.

Section 2.4

1. $y(x) = Ae^{-3x} + Be^{-x} + \frac{1}{3}x - \frac{1}{9}$

3. $y(x) = e^{-x}[A\cos(x) + B\sin(x)] + x^2 - x + 2$

5. $y(x) = A + Be^{-2x} + \frac{1}{2}x^2 + 2x + \frac{1}{2}e^{-2x}$

7. $y(x) = (A + Bx)e^{-2x} + \left(\frac{1}{9}x - \frac{2}{27}\right) e^{x}$

9. $y(x) = A\cos(3x) + B\sin(3x) + \frac{1}{12}x^2 \sin(3x) + \frac{1}{36}x\cos(3x)$

11. $y(x) = \dfrac{2ax - 1}{8a^2} + \dfrac{\omega^2 e^{-2ax}}{8a^2(a^2 + \omega^2)} - \dfrac{a\sin(2\omega x) - \omega\cos(2\omega x)}{8\omega(a^2 + \omega^2)}$

Section 2.5

1. $\gamma = 3$,

5. $x(t) = e^{-ct/(2m)} [A\cos(\omega_0 t) + B\sin(\omega_0 t)] + \dfrac{F_0 \sin(\omega t - \varphi)}{\sqrt{c^2\omega^2 + (k - m\omega^2)^2}}$

Section 2.6

1. $y(x) = Ae^{x} + Be^{3x} + \frac{1}{8}e^{-x}$

3. $y(x) = Ae^{2x} + Be^{-2x} - (3x + 2)e^{x}/9$

5. $y(x) = (A + Bx)e^{-2x} + x^3 e^{-2x}/6$

7. $y(x) = Ae^{2x} + Bxe^{2x} + \left(\frac{1}{2}x^2 + \frac{1}{6}x^3\right) e^{2x}$

9. $y(x) = Ae^{x} + Bxe^{x} + x\ln(x)e^{x}$

Section 2.7

1. $y(x) = C_1 x + C_2 x^{-1}$

3. $y(x) = C_1 x^2 + C_2/x$

5. $y(x) = C_1/x + C_2 \ln(x)/x$

7. $y(x) = C_1 x\cos[2\ln(x)] + C_2 x\sin[\ln(x)]$

9. $y(x) = C_1 \cos[\ln(x)] + C_2 \sin[\ln(x)]$

11. $y(x) = C_1 x^2 + C_2 x^4 + C_3/x$

Section 2.8

1. The trajectories spiral outward from $(0, 0)$.

3. The equilibrium points are $(x, 0)$; they are unstable.

5. The equilibrium points are $v = 0$ and $|x| < 2$; they are unstable.

Section 3.1

1. $A + B = \begin{pmatrix} 4 & 5 \\ 3 & 4 \end{pmatrix} = B + A$

3. $3A - 2B = \begin{pmatrix} 7 & 10 \\ -1 & 2 \end{pmatrix}, \qquad 3(2A - B) = \begin{pmatrix} 15 & 21 \\ 0 & 6 \end{pmatrix}$

5. $(A + B)^T = \begin{pmatrix} 4 & 3 \\ 5 & 4 \end{pmatrix}, \qquad A^T + B^T = \begin{pmatrix} 4 & 3 \\ 5 & 4 \end{pmatrix}$

7. $AB = \begin{pmatrix} 11 & 11 \\ 5 & 5 \end{pmatrix}, \; A^T B = \begin{pmatrix} 5 & 5 \\ 8 & 8 \end{pmatrix}, \; BA = \begin{pmatrix} 4 & 6 \\ 8 & 12 \end{pmatrix}, \; B^T A = \begin{pmatrix} 5 & 8 \\ 5 & 8 \end{pmatrix}$

9. $BB^T = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}, \qquad B^T B = \begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix}$

11. $A^3 + 2A = \begin{pmatrix} 65 & 100 \\ 25 & 40 \end{pmatrix}$

13. yes $\begin{pmatrix} 27 & 11 \\ 2 & 5 \end{pmatrix}$

15. yes $\begin{pmatrix} 11 & 8 \\ 8 & 4 \\ 5 & 3 \end{pmatrix}$

17. no

19. $5(2A) = \begin{pmatrix} 10 & 10 \\ 10 & 20 \\ 30 & 10 \end{pmatrix} = 10A$

21. $(A + B) + C = \begin{pmatrix} 4 & 0 \\ 8 & 2 \end{pmatrix} = A + (B + C)$

23. $A(B + C) = \begin{pmatrix} 9 & -1 \\ 11 & -2 \end{pmatrix} = AB + AC$

27. $\begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$

29. $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 3 & 0 & -4 & -4 \\ 1 & 1 & 1 & 1 \\ 2 & -3 & 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -3 \\ 7 \end{pmatrix}$

Section 3.2

1. 7

3. 1

5. -24

7. 3

Section 3.3

1. $x_1 = \frac{9}{5}$, $x_2 = \frac{3}{5}$

3. $x_1 = 0$, $x_2 = 0$, $x_3 = -2$

Section 3.4

1. $x_1 = 1, x_2 = 2$

3. $x_1 = x_3 = \alpha, x_2 = -\alpha$

5. $x_1 = -1, x_2 = 2\alpha, x_3 = \alpha$

7. $x_1 = 1, x_2 = 2.6, x_3 = 2.2$

Section 3.5

1. $\lambda = 4, \quad \mathbf{x}_0 = \alpha \begin{pmatrix} 2 \\ 1 \end{pmatrix}; \qquad \lambda = -3 \quad \mathbf{x}_0 = \beta \begin{pmatrix} 1 \\ -3 \end{pmatrix}$

3. $\lambda = 1 \quad \mathbf{x}_0 = \alpha \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}; \qquad \lambda = 0, \quad \mathbf{x}_0 = \gamma \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

5. $\lambda = 1, \quad \mathbf{x}_0 = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}; \qquad \lambda = 2, \quad \mathbf{x}_0 = \gamma \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

7. $\lambda = 0, \quad \mathbf{x}_0 = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}; \ \lambda = 1, \quad \mathbf{x}_0 = \beta \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}; \ \lambda = 2, \quad \mathbf{x}_0 = \gamma \begin{pmatrix} 7 \\ 3 \\ 1 \end{pmatrix}$

Section 3.6

1. $\mathbf{x} = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-t} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{3t}$

3. $\mathbf{x} = c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{3t} + c_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^{-t}$

5. $\mathbf{x} = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{t/2} + c_2 \begin{pmatrix} t \\ -1/2 - t \end{pmatrix} e^{t/2}$

7. $\mathbf{x} = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{2t} + c_2 \begin{pmatrix} -1 + t \\ -t \end{pmatrix} e^{2t}$

9. $\mathbf{x} = c_3 \begin{pmatrix} -3\cos(2t) - 2\sin(2t) \\ \cos(2t) \end{pmatrix} e^{t} + c_4 \begin{pmatrix} 2\cos(2t) - 3\sin(2t) \\ \sin(2t) \end{pmatrix} e^{t}$

11. $\mathbf{x} = c_3 \begin{pmatrix} 2\cos(t) \\ 7\cos(t) + \sin(t) \end{pmatrix} e^{-3t} + c_4 \begin{pmatrix} 2\sin(t) \\ 7\sin(t) - \cos(t) \end{pmatrix} e^{-3t}$

13. $\mathbf{x} = c_3 \begin{pmatrix} -\cos(2t) + \sin(2t) \\ \cos(2t) \end{pmatrix} e^{t} + c_4 \begin{pmatrix} -\cos(2t) - \sin(2t) \\ \sin(2t) \end{pmatrix} e^{t}$

15. $\mathbf{x} = c_1 \begin{pmatrix} -1 \\ 2 \end{pmatrix} e^{3t} + c_2 \begin{pmatrix} -3 \\ 2 \end{pmatrix} e^{-t}$

17. $\mathbf{x} = c_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} e^{t} + c_3 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} e^{2t}$

19. $\mathbf{x} = c_1 \begin{pmatrix} 3 \\ -2 \\ 12 \end{pmatrix} e^{-t} + c_2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} e^{t} + c_3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} e^{2t}$

Section 3.7

1. $\begin{pmatrix} e^t & 3te^t \\ 0 & e^t \end{pmatrix}$ 3. $\begin{pmatrix} e^t & te^t & \frac{1}{2}t^2e^t \\ 0 & e^t & te^t \\ 0 & 0 & e^t \end{pmatrix}$ 5. $\begin{pmatrix} e^t & 2te^t & 2t^2e^t \\ 0 & e^t & 2te^t \\ 0 & 0 & e^t \end{pmatrix}$

7.
$$x_1(t) = \tfrac{1}{2}\left(3e^t - e^{-t}\right)x_1(0) + \tfrac{1}{2}\left(e^{-t} - e^t\right)x_2(0) + 3te^t - 3\sinh(t) + 2t$$
$$x_2(t) = \tfrac{1}{2}\left(3e^t - 3e^{-t}\right)x_1(0) + \tfrac{1}{2}\left(3e^{-t} - e^t\right)x_2(0) + 3te^t + 2e^{-t} - 5\sinh(t) + 4t - 2$$

9.
$$x_1(t) = [\cos(t) + 2\sin(t)]x_1(0) - 5\sin(t)x_2(0) + 2t\cos(t) - t\sin(t) - 3\sin(t)$$
$$x_2(t) = \sin(t)x_1(0) + [\cos(t) - 2\sin(t)]x_2(0) + t\cos(t) - \sin(t)$$

11.
$$\begin{aligned} x_1(t) &= \tfrac{1}{3}\left[e^{4t} + 2e^t\right]x_1(0) + \tfrac{1}{3}\left[e^{4t} - e^t\right]x_2(0) + \tfrac{1}{3}\left[e^{4t} - e^t\right]x_3(0) \\ &\quad - \frac{t^2}{6}e^t - \frac{4t}{9}e^t + \frac{4}{27}e^{4t} - \frac{4}{27}e^t \\ x_2(t) &= \tfrac{1}{3}\left[e^{4t} - e^t\right]x_1(0) + \tfrac{1}{3}\left[e^{4t} + 2e^t\right]x_2(0) + \tfrac{1}{3}\left[e^{4t} - e^t\right]x_3(0) \\ &\quad + \frac{t^2}{3}e^t - \frac{4t}{9}e^t + \frac{4}{27}e^{4t} - \frac{4}{27}e^t \\ x_3(t) &= \tfrac{1}{3}\left[e^{4t} - e^t\right]x_1(0) + \tfrac{1}{3}\left[e^{4t} - e^t\right]x_2(0) + \tfrac{1}{3}\left[e^{4t} + 2e^t\right]x_3(0) \\ &\quad - \frac{t^2}{6}e^t + \frac{5t}{6}e^t + \frac{4}{27}e^{4t} - \frac{4}{27}e^t \end{aligned}$$

13.
$$x_1(t) = \left(\tfrac{1}{2}e^{3t} + \tfrac{1}{2}e^{-t}\right)x_1(0) + \left(\tfrac{1}{4}e^{3t} - \tfrac{1}{4}e^{-t}\right)x_3(0) + \tfrac{5}{4}\left(e^t - e^{3t}\right) + \tfrac{1}{4}\left(e^{-t} - e^t\right)$$
$$x_2(t) = e^{-2t}x_2(0) + 2\left(e^t - e^{-2t}\right)$$
$$x_3(t) = \left(e^{3t} - e^{-t}\right)x_1(0) + \left(\tfrac{1}{2}e^{3t} + \tfrac{1}{2}e^{-t}\right)x_3(0) + \tfrac{5}{2}\left(e^t - e^{3t}\right) + \tfrac{1}{2}\left(e^t - e^{-t}\right)$$

15.
$$x_1(t) = (1-t)e^{2t}x_1(0) + te^{2t}x_2(0) + (2t + t^2)e^{2t}x_3(0) + \left(t^2 + \tfrac{1}{4}t + \tfrac{1}{4}\right)e^{2t} - \tfrac{3}{4}t - \tfrac{1}{4}$$
$$\begin{aligned} x_2(t) = &-te^{2t}x_1(0) + (1+t)e^{2t}x_2(0) + (4t + t^2)e^{2t}x_3(0) \\ &+ \left(t^2 + \tfrac{9}{4}t - \tfrac{3}{2}\right)e^{2t} + 2e^t - \tfrac{1}{4}t - \tfrac{1}{2} \end{aligned}$$
$$x_3(t) = e^{2t}x_3(0) + e^{2t} - e^t$$

Section 4.1

1. $\mathbf{a} \times \mathbf{b} = -3\mathbf{i} + 19\mathbf{j} + 10\mathbf{k}$ 3. $\mathbf{a} \times \mathbf{b} = \mathbf{i} - 8\mathbf{j} + 7\mathbf{k}$ 5. $\mathbf{a} \times \mathbf{b} = -3\mathbf{i} - 2\mathbf{j} - 5\mathbf{k}$

9. $\nabla f = y\cos(yz)\mathbf{i} + [x\cos(yz) - xyz\sin(yz)]\mathbf{j} - xy^2\sin(yz)\mathbf{k}$

11. $\nabla f = 2xy^2(2z+1)^2\mathbf{i} + 2x^2y(2z+1)^2\mathbf{j} + 4x^2y^2(2z+1)\mathbf{k}$

13. Plane parallel to the xy plane at height of $z = 3$, $\mathbf{n} = \mathbf{k}$

15. Paraboloid,
$$\mathbf{n} = -\frac{2x\,\mathbf{i}}{\sqrt{1+4x^2+4y^2}} - \frac{2y\,\mathbf{j}}{\sqrt{1+4x^2+4y^2}} + \frac{\mathbf{k}}{\sqrt{1+4x^2+4y^2}}$$

17. A plane, $\mathbf{n} = \mathbf{j}/\sqrt{2} - \mathbf{k}/\sqrt{2}$

19. A parabola of infinite extent along the y-axis, $\mathbf{n} = -2x\mathbf{i}/\sqrt{1+4x^2} + \mathbf{k}/\sqrt{1+4x^2}$

21. $y = 2/(x+1); \quad z = \exp[(y-1)/y]$

23. $y = x; \quad z^2 = y/(3y-2)$

Section 4.2

1. $\nabla\cdot\mathbf{F} = 2xz + z^2, \quad \nabla\times\mathbf{F} = (2xy - 2yz)\mathbf{i} + (x^2 - y^2)\mathbf{j}, \quad \nabla(\nabla\cdot\mathbf{F}) = 2z\mathbf{i} + (2x+2z)\mathbf{k}$

3. $\nabla\cdot\mathbf{F} = 2(x-y) - xe^{-xy} + xe^{2y}, \quad \nabla\times\mathbf{F} = 2xze^{2y}\mathbf{i} - ze^{2y}\mathbf{j} + [2(x-y) - ye^{-xy}]\,\mathbf{k},$
$\nabla(\nabla\cdot\mathbf{F}) = \left(2 - e^{-xy} + xye^{-xy} + e^{2y}\right)\mathbf{i} + \left(x^2e^{-xy} + 2xe^{2y} - 2\right)\mathbf{j}$

5. $\nabla\cdot\mathbf{F} = 0, \quad \nabla\times\mathbf{F} = -x^2\mathbf{i} + (5y - 9x^2)\mathbf{j} + (2xz - 5z)\mathbf{k}, \quad \nabla(\nabla\cdot\mathbf{F}) = \mathbf{0}$

7. $\nabla\cdot\mathbf{F} = e^{-y} + z^2 - 3e^{-z}, \ \nabla\times\mathbf{F} = -2yz\mathbf{i} + xe^{-y}\mathbf{k}, \ \nabla(\nabla\cdot\mathbf{F}) = -e^{-y}\mathbf{j} + (2z + 3e^{-z})\mathbf{k}$

9. $\nabla\cdot\mathbf{F} = yz + x^3ze^z + xye^z,$
$\nabla\times\mathbf{F} = (xe^z - x^3ye^z - x^3yze^z)\mathbf{i} + (xy - ye^z)\mathbf{j} + (3x^2yze^z - xz)\mathbf{k},$
$\nabla(\nabla\cdot\mathbf{F}) = \left(3x^2ze^z + ye^z\right)\mathbf{i} + \left(z + xe^z\right)\mathbf{j} + \left(y + x^3e^z + x^3ze^z + xye^z\right)\mathbf{k},$

11. $\nabla\cdot\mathbf{F} = y^2 + xz^2 - xy\sin(z),$
$\nabla\times\mathbf{F} = [x\cos(z) - 2xyz]\mathbf{i} - y\cos(z)\mathbf{j} + (yz^2 - 2xy)\mathbf{k},$
$\nabla(\nabla\cdot\mathbf{F}) = [z^2 - y\sin(z)]\mathbf{i} + \big[2y - x\sin(z)\big]\mathbf{j} + \big[2xz - xy\cos(z)\big]\mathbf{k}$

13. $\nabla\cdot\mathbf{F} = y^2 + xz - xy\sin(z),$
$\nabla\times\mathbf{F} = [x\cos(z) - xy]\mathbf{i} - y\cos(z)\mathbf{j} + (yz - 2xy)\mathbf{k},$
$\nabla(\nabla\cdot\mathbf{F}) = [z - y\sin(z)]\mathbf{i} + \big[2y - x\sin(z)\big]\mathbf{j} + \big[x - xy\cos(z)\big]\mathbf{k}$

Section 4.3

1. $16/7 + 2/(3\pi)$ 3. $e^2 + 2e^8/3 + e^{64}/2 - 13/6$ 5. -4π 7. 0 9. 2π

Section 4.4

1. $\varphi(x,y,z) = x^2y + y^2z + 4z + \text{constant}$ 3. $\varphi(x,y,z) = xyz + \text{constant}$

5. $\varphi(x,y,z) = x^2\sin(y) + xe^{3z} + 4z + \text{constant}$

7. $\varphi(x,y,z) = xe^{2z} + y^3 + \text{constant}$ 9. $\varphi(x,y,z) = xy + xz + \text{constant}$

Section 4.5

1. 1/2 3. 0 5. 27/2 7. 5

9. 0 11. 40/3 13. 86/3 15. 96π

Section 4.6

1. -5 3. 1 5. 0 7. 0 9. -16π 11. -2

Section 4.7

1. -10 3. 2 5. π 7. $45/2$

Section 4.8

1. 3 3. -16 5. 4π 7. $5/12$

Section 5.1

1. $f(t) = \frac{1}{2} - \frac{2}{\pi}\sum_{m=1}^{\infty}\frac{\sin[(2m-1)t]}{2m-1}$

3. $f(t) = -\frac{\pi}{4} + \sum_{n=1}^{\infty}\frac{(-1)^n - 1}{n^2\pi}\cos(nt) + \frac{1-2(-1)^n}{n}\sin(nt),$

5. $f(t) = \frac{\pi}{8} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{2\cos(n\pi/2)\sin^2(n\pi/4)}{n^2}\cos(nt) + \frac{\sin(n\pi/2)}{n^2}\sin(nt)$

7. $$f(t) = \frac{\sinh(aL)}{aL} + 2aL\sinh(aL)\sum_{n=1}^{\infty}\frac{(-1)^n}{a^2L^2+n^2\pi^2}\cos\left(\frac{n\pi t}{L}\right) - 2\pi\sinh(aL)\sum_{n=1}^{\infty}\frac{n(-1)^n}{a^2L^2+n^2\pi^2}\sin\left(\frac{n\pi t}{L}\right)$$

9. $f(t) = \frac{1}{\pi} + \frac{1}{2}\sin(t) - \frac{2}{\pi}\sum_{m=1}^{\infty}\frac{\cos(2mt)}{4m^2-1}$

11. $f(t) = \frac{a}{2} - \frac{4a}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^2}\cos\left[\frac{(2m-1)\pi t}{a}\right] - \frac{2a}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin\left(\frac{n\pi t}{a}\right)$

13. $f(t) = \frac{\pi-1}{2} + \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\sin(n\pi t)}{n}$, 15. $f(t) = \frac{4a\cosh(a\pi/2)}{\pi}\sum_{m=1}^{\infty}\frac{\cos[(2m-1)t]}{a^2+(2m-1)^2}$

Section 5.3

1. $f(x) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{m=1}^{\infty}\frac{\cos[(2m-1)x]}{(2m-1)^2}$, $f(x) = \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}\sin(nx)}{n}$

3. $f(x) = \frac{a^3}{6} - \frac{a^2}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{m^2}\cos\left(\frac{2m\pi x}{a}\right)$, $f(x) = \frac{8a^2}{\pi^3}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^3}\sin\left[\frac{(2m-1)\pi x}{a}\right]$

5. $f(x) = \frac{1}{4} - \frac{2}{\pi^2}\sum_{m=1}^{\infty}\frac{\cos[2(2m-1)\pi x]}{(2m-1)^2}$, $f(x) = \frac{4}{\pi^2}\sum_{m=1}^{\infty}\frac{(-1)^{m+1}\sin[(2m-1)\pi x]}{(2m-1)^2}$

7. $f(x) = \frac{2\pi^2}{3} - 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx)$, $f(x) = 2\pi\sum_{n=1}^{\infty}\frac{\sin(nx)}{n} + \frac{8}{\pi}\sum_{m=1}^{\infty}\frac{\sin[(2m-1)x]}{(2m-1)^3}$

9. $f(x) = \frac{a}{6} + \frac{4a}{\pi^2}\sum_{m=1}^{\infty} \frac{(-1)^m \sin[(2m-1)\pi/6]}{(2m-1)^2} \cos\left[\frac{(2m-1)\pi x}{a}\right]$

$$f(x) = \frac{a}{\pi^2}\sum_{m=1}^{\infty} \frac{(-1)^m \sin(m\pi/3)}{m^2} \sin\left(\frac{2m\pi x}{a}\right) - \frac{2a}{3\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin\left(\frac{n\pi x}{a}\right)$$

11. $f(x) = \frac{3}{4} + \frac{1}{\pi}\sum_{m=1}^{\infty} \frac{(-1)^m}{2m-1} \cos\left[\frac{(2m-1)\pi x}{a}\right]$

$$f(x) = \frac{1}{\pi}\sum_{n=1}^{\infty} \frac{1+\cos(n\pi/2) - 2(-1)^n}{n} \sin\left(\frac{n\pi x}{a}\right)$$

13. $f(x) = \frac{3a}{8} + \frac{2a}{\pi^2}\sum_{n=1}^{\infty} \frac{\cos(n\pi/2) - 1}{n^2} \cos\left(\frac{n\pi x}{a}\right),$

$$f(x) = \frac{a}{\pi}\sum_{n=1}^{\infty}\left[\frac{2}{n^2\pi}\sin\left(\frac{n\pi}{2}\right) - \frac{(-1)^n}{n}\right]\sin\left(\frac{n\pi x}{a}\right)$$

Section 5.4

1. $f(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(2n-1)t]}{2n-1}, \quad f(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)t - \pi/2]}{2n-1}$

3. $f(t) = 2\sum_{n=1}^{\infty}\frac{1}{n}\cos\left[nt + (-1)^n\frac{\pi}{2}\right], \quad f(t) = 2\sum_{n=1}^{\infty}\frac{1}{n}\sin\left\{nt + [1+(-1)^n]\frac{\pi}{2}\right\}$

Section 5.5

1. $f(t) = \frac{\pi}{2} - \frac{2}{\pi}\sum_{m=-\infty}^{\infty}\frac{e^{i(2m-1)t}}{(2m-1)^2}$

3. $f(t) = 1 + \frac{i}{\pi}\sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty}\frac{e^{n\pi i t}}{n}$

5. $f(t) = \frac{1}{2} - \frac{i}{\pi}\sum_{m=-\infty}^{\infty}\frac{e^{2(2m-1)it}}{2m-1}$

Section 5.6

1. $y(t) = A\cosh(t) + B\sinh(t) - \frac{1}{2} - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin[(2n-1)t]}{(2n-1)+(2n-1)^3}$

3. $y(t) = Ae^{2t} + Be^{t} + \frac{1}{4} + \frac{6}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)t]}{[2-(2n-1)^2]^2 + 9(2n-1)^2}$

$$+\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{[2-(2n-1)^2]\sin[(2n-1)t]}{(2n-1)\{[2-(2n-1)^2]^2+9(2n-1)^2\}}$$

5. $y_p(t) = \frac{\pi}{8} - \frac{2}{\pi}\sum_{n=-\infty}^{\infty}\frac{e^{i(2n-1)t}}{(2n-1)^2[4-(2n-1)^2]}$

7. $q(t) = \sum_{n=-\infty}^{\infty}\frac{\omega^2\varphi_n}{(in\omega_0)^2 + 2i\alpha n\omega_0 + \omega^2}e^{in\omega_0 t}$

Section 5.7

1. $f(t) = \frac{3}{2} - \cos(\pi x/2) - \sin(\pi x/2) - \frac{1}{2}\cos(\pi x)$

Section 6.1

1. $\lambda_n = (2n-1)^2\pi^2/(4L^2)$, $y_n(x) = \cos[(2n-1)\pi x/(2L)]$

3. $\lambda_0 = -1$, $y_0(x) = e^{-x}$ and $\lambda_n = n^2$, $y_n(x) = \sin(nx) - n\cos(nx)$

5. $\lambda_n = -n^4\pi^4/L^4$, $y_n(x) = \sin(n\pi x/L)$

7. $\lambda_n = k_n^2$, $y_n(x) = \sin(k_n x)$ with $k_n = -\tan(k_n)$

9. $\lambda_0 = -m_0^2$, $y_0(x) = \sinh(m_0 x) - m_0\cosh(m_0 x)$ with $\coth(m_0\pi) = m_0$; $\lambda_n = k_n^2$, $y_n(x) = \sin(k_n x) - k_n\cos(k_n x)$ with $k_n = -\cot(k_n\pi)$

11.

(a) $\lambda_n = n^2\pi^2$, $\quad y_n(x) = \sin[n\pi\ln(x)]$

(b) $\lambda_n = (2n-1)^2\pi^2/4$, $\quad y_n(x) = \sin[(2n-1)\pi\ln(x)/2]$

(c) $\lambda_0 = 0$, $\quad y_0(x) = 1$; $\quad \lambda_n = n^2\pi^2$, $\quad y_n(x) = \cos[n\pi\ln(x)]$

13. $\lambda_n = n^2+1$, $y_n(x) = \sin[n\ln(x)]/x$

15. $\lambda = 0$, $y_0(x) = 1$; $y_n(x) = \cosh(\lambda_n x) + \cos(\lambda_n x) - \tanh(\lambda_n)[\sinh(\lambda_n x) + \sin(\lambda_n x)]$, where $n = 1, 2, 3, \ldots$, and λ_n is the nth root of $\tanh(\lambda) = -\tan(\lambda)$.

Section 6.3

1. $f(x) = \dfrac{2}{\pi}\sum_{n=1}^{\infty}\dfrac{(-1)^{n+1}}{n}\sin\left(\dfrac{n\pi x}{L}\right)$ $\qquad$ 3. $f(x) = \dfrac{8L}{\pi^2}\sum_{n=1}^{\infty}\dfrac{(-1)^{n+1}}{(2n-1)^2}\sin\left[\dfrac{(2n-1)\pi x}{2L}\right]$

Section 6.4

1. $f(x) = \frac{1}{4}P_0(x) + \frac{1}{2}P_1(x) + \frac{5}{16}P_2(x) + \cdots$ $\quad$ 3. $f(x) = \frac{1}{2}P_0(x) + \frac{5}{8}P_2(x) - \frac{3}{16}P_4(x) + \cdots$
5. $f(x) = \frac{3}{2}P_1(x) - \frac{7}{8}P_3(x) + \frac{11}{16}P_5(x) + \cdots$

Section 7.3

1. $u(x,t) = \dfrac{4L}{c\pi^2}\sum_{m=1}^{\infty}\dfrac{1}{(2m-1)^2}\sin\left[\dfrac{(2m-1)\pi x}{L}\right]\sin\left[\dfrac{(2m-1)\pi ct}{L}\right]$

3. $u(x,t) = \dfrac{9h}{\pi^2}\sum_{n=1}^{\infty}\dfrac{1}{n^2}\sin\left(\dfrac{2n\pi}{3}\right)\sin\left(\dfrac{n\pi x}{L}\right)\cos\left(\dfrac{n\pi ct}{L}\right)$

5. $u(x,t) = \sin\left(\dfrac{\pi x}{L}\right)\cos\left(\dfrac{\pi ct}{L}\right)$
$$+\frac{4aL}{\pi^2 c}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{(2n-1)^2}\sin\left[\frac{(2n-1)\pi}{4}\right]\sin\left[\frac{(2n-1)\pi x}{L}\right]\sin\left[\frac{(2n-1)\pi ct}{L}\right]$$

7. $u(x,t) = \dfrac{4L}{\pi^2}\sum_{n=1}^{\infty}\dfrac{(-1)^{n+1}}{(2n-1)^2}\sin\left[\dfrac{(2n-1)\pi x}{L}\right]\cos\left[\dfrac{(2n-1)\pi ct}{L}\right]$

9. $u(x,t) = \sqrt{a}\sum_{n=1}^{\infty} \frac{J_1(2k_n\sqrt{a}\,)J_0(2k_n\sqrt{x}\,)}{k_n^2 J_1^2(2k_n)} \sin(k_n t),$
where k_n is the nth solution of $J_0(2k) = 0$.

11. $u(x,t) = \frac{8Lcs_0}{\pi^2}\sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)^2} \cos\left[\frac{(2n-1)\pi x}{2L}\right] \sin\left[\frac{(2n-1)\pi ct}{2L}\right]$

Section 7.4

1. $u(x,t) = \sin(2x)\cos(2ct) + \cos(x)\sin(ct)/c$

3. $u(x,t) = \frac{1+x^2+c^2t^2}{(1+x^2+c^2t^2)^2+4x^2c^2t^2} + \frac{e^x\sinh(ct)}{c}$

5. $u(x,t) = \cos\left(\frac{\pi x}{2}\right)\cos\left(\frac{\pi ct}{2}\right) + \frac{\sinh(ax)\sinh(act)}{ac}$

Section 8.3

1. $u(x,t) = \frac{4A}{\pi}\sum_{m=1}^{\infty} \frac{\sin[(2m-1)x]}{2m-1} e^{-a^2(2m-1)^2t}$

3. $u(x,t) = -2\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin(nx)e^{-a^2n^2t}$

5. $u(x,t) = \frac{4}{\pi}\sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{(2m-1)^2} \sin[(2m-1)x]e^{-a^2(2m-1)^2t}$

7. $u(x,t) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{m=1}^{\infty} \frac{\cos[(2m-1)x]}{(2m-1)^2} e^{-a^2(2m-1)^2t}$

9. $u(x,t) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty} \frac{\cos[(2n-1)x]}{(2n-1)^2} e^{-a^2(2n-1)^2t}$

11. $u(x,t) = \frac{32}{\pi}\sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)^3} \cos\left[\frac{(2n-1)x}{2}\right] e^{-a^2(2n-1)^2t/4}$

13. $u(x,t) = \frac{4}{\pi}\sum_{n=1}^{\infty} \frac{\sin[(2n-1)x/2]}{2n-1} e^{-a^2(2n-1)^2t/4}$

15. $u(x,t) = \sum_{n=1}^{\infty} \left[\frac{4}{2n-1} - \frac{8(-1)^{n+1}}{(2n-1)^2\pi}\right] \sin\left[\frac{(2n-1)x}{2}\right] e^{-a^2(2n-1)^2t/4}$

17. $u(x,t) = \frac{T_0x}{\pi} + \frac{2T_0}{\pi}\sum_{n=1}^{\infty} \frac{1}{n}\sin(nx)e^{-a^2n^2t}$

19. $u(x,t) = h_1 + \frac{(h_2-h_1)x}{L} + \frac{2(h_2-h_1)}{\pi}\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\left(\frac{n\pi x}{L}\right)\exp\left(-\frac{a^2n^2\pi^2}{L^2}t\right)$

21. $u(x,t) = h_0 - \frac{4h_0}{\pi}\sum_{n=1}^{\infty} \frac{1}{2n-1} \sin\left[\frac{(2n-1)\pi x}{L}\right] \exp\left[-\frac{(2n-1)^2\pi^2a^2t}{L^2}\right]$

23. $u(x,t) = \frac{1}{3} - t - \frac{2}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(n\pi x) e^{-a^2 n^2 \pi^2 t}$

25. $u(x,t) = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)^4} \sin[(2n-1)x] \left[1 - e^{-(2n-1)^2 t}\right]$

27. $u(x,t) = \frac{A_0(L^2 - x^2)}{2\kappa} + \frac{A_0 L}{h}$
$$- \frac{2L^2 A_0}{\kappa} \sum_{n=1}^{\infty} \frac{\sin(\beta_n)}{\beta_n^4 [1 + \kappa \sin^2(\kappa)/hL]} \cos\left(\frac{\beta_n x}{L}\right) \exp\left(-\frac{a^2 \beta_n^2 t}{L^2}\right),$$
where β_n is the nth root of $\beta \tan(\beta) = hL/\kappa$.

29. $u(x,t) = 4u_0 \sum_{n=1}^{\infty} \frac{\sin(k_n L)\cos(k_n x)}{2k_n L + \sin(2k_n L)} \cos(k_n x) \exp\left[-(k_1 + a^2 k_n^2) t\right]$,
where k_n denotes the nth root of $k \tan(kL) = k_2/a^2$

31. $u(r,t) = \frac{2}{\pi r} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(n\pi r) e^{-a^2 n^2 \pi^2 t}$

33. $u(r,t) = \frac{4bu_0}{r} \sum_{n=1}^{\infty} \frac{\sin(k_n) - k_n \cos(k_n)}{k_n [2k_n - \sin(2k_n)]} \sin\left(\frac{k_n r}{b}\right) e^{-a^2 k_n^2 t/b^2}$,
where k_n is the nth root of $k \cot(k) = 1 - A$, $n\pi < k_n < (n+1)\pi$.

35. $u(r,t) = \theta + 2(1-\theta) \sum_{n=1}^{\infty} \frac{J_0(k_n r/b)}{k_n J_1(k_n)} e^{-a^2 k_n^2 t/b^2}$,
where k_n is the nth root of $J_0(k) = 0$.

37. $u(r,t) = \frac{G}{4\rho\nu}(b^2 - r^2) - \frac{2Gb^2}{\rho\nu} \sum_{n=1}^{\infty} \frac{J_0(k_n r/b)}{k_n^3 J_1(k_n)} e^{-\nu k_n^2 t/b^2}$,
where k_n is the nth root of $J_0(k) = 0$.

39. $u(r,t) = \frac{2T_0}{L^2} e^{-\kappa t} \sum_{n=1}^{\infty} \frac{[LJ_1(k_n L) - bJ_1(k_n b)] J_0(k_n r)}{k_n [J_0^2(k_n L) + J_1^2(k_n L)]} e^{-a^2 k_n^2 t}$,
where k_n is the nth root of $kJ_1(kL) = hJ_0(kL)$.

Section 9.3

1. $u(x,y) = \frac{4}{\pi} \sum_{m=1}^{\infty} \frac{\sinh[(2m-1)\pi(a-x)/b] \sin[(2m-1)\pi y/b]}{(2m-1)\sinh[(2m-1)\pi a/b]}$

3. $u(x,y) = -\frac{2a}{\pi} \sum_{n=1}^{\infty} \frac{\sinh(n\pi y/a)\sin(n\pi x/a)}{n \ \sinh(n\pi b/a)}$

5. $u(x,y) = \frac{4}{\pi} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sinh[(2n-1)\pi y/2a] \cos[(2n-1)\pi x/2a]}{(2n-1)\sinh[(2n-1)\pi b/2a]}$

7. $u(x,y) = 1$

9. $u(x,y) = 1 - \frac{4}{\pi} \sum_{m=1}^{\infty} \frac{\cosh[(2m-1)\pi y/a] \sin[(2m-1)\pi x/a]}{(2m-1)\cosh[(2m-1)\pi b/a]}$

11. $u(x,y) = 1$

13. $u(x,y) = T_0 + \Delta T \cos(2\pi x/\lambda)e^{-2\pi y/\lambda}$

15. $u(x,y) = L - y - \dfrac{4\gamma L}{\pi^2} \displaystyle\sum_{m=1}^{\infty} \dfrac{\cosh[(2m-1)\pi x/L]}{(2m-1)^2 \sinh[(2m-1)\pi/L]} \sin\left(\dfrac{n\pi y}{L}\right)$

17. $u(r,z) = \dfrac{Aa^2 z}{b^2} + \dfrac{2Aa}{b^2} \displaystyle\sum_{n=1}^{\infty} \dfrac{\sinh(k_n z)\, J_1(k_n a)\, J_0(k_n r)}{k_n^2 \cosh(k_n L) J_0^2(k_n b)}$,
where k_n is nth positive zero of $J_0'(kb) = -J_1(kb) = 0$.

19. $u(r,z) = \dfrac{(z-h)r_0^2}{a^2} + 2r_0 \displaystyle\sum_{n=1}^{\infty} \dfrac{\sinh[k_n(z-h)/a] J_1(k_n r_0/a) J_0(k_n r/a)}{k_n^2 \cosh(k_n h/a) J_0^2(k_n)}$,
where k_n is nth positive zero of $J_1(k) = 0$.

21. Case (a):
$u(r,z) = \dfrac{4}{\pi} \displaystyle\sum_{n=1}^{\infty} \dfrac{I_1[(2n-1)\pi r/2]}{I_1[(2n-1)\pi/2]} \dfrac{\sin[(2n-1)\pi(z-1)/2]}{2n-1}$
Case (b):
$u(r,z) = -2 \displaystyle\sum_{n=1}^{\infty} \dfrac{\cosh(k_n z)}{\cosh(k_n)} \dfrac{J_1(k_n r)}{k_n J_0(k_n)}$
where k_n is the nth root of $J_1(k) = 0$.

23. $u(r,z) = -\dfrac{2}{\pi} \displaystyle\sum_{n=1}^{\infty} \dfrac{(-1)^n I_1(n\pi r) \sin(n\pi z)}{n\, I_1(n\pi a)}$

25. $u(r,z) = 2 \displaystyle\sum_{n=1}^{\infty} \dfrac{\sinh(k_n z) J_1(k_n r)}{k_n^3 \cosh(k_n a) J_1(k_n)}$,
where k_n is nth positive zero of $kJ_0(k) = J_1(k)$.

27. $u(r,z) = 2u_0 \displaystyle\sum_{n=1}^{\infty} \dfrac{J_1(\mu_n) J_0(\mu_n r/a)}{\mu_n[J_0^2(\mu_n) + J_1^2(\mu_n)]} \dfrac{\cosh[\mu_n(L-z)/a]}{\cosh(\mu_n L/a)}$,
where $\mu_n J_1(\mu_n) = \beta J_0(\mu_n)$, $\beta = aK/D$ and $n = 1, 2, 3, \ldots$.

29. $u(r,z) = -V \left\{ 1 - 2 \displaystyle\sum_{n=1}^{\infty} \dfrac{\cosh(k_n z/a)}{k_n\, J_1(k_n)} \exp\left(-\dfrac{k_n d}{2a}\right) J_0\left(\dfrac{k_n r}{a}\right) \right\}$
if $|z| < d/2$, and
$u(r,z) = -2V \displaystyle\sum_{n=1}^{\infty} \dfrac{\sinh[k_n d/(2a)]}{k_n\, J_1(k_n)} \exp\left(-\dfrac{k_n |z|}{a}\right) J_0\left(\dfrac{k_n r}{a}\right)$
if $|z| > d/2$.

31. $u(r,z) = 2B \displaystyle\sum_{n=1}^{\infty} \dfrac{\exp[z(1-\sqrt{1+4k_n^2}\,)/2] J_0(k_n r)}{(k_n^2 + B^2) J_0(k_n)}$,
where k_n is nth positive zero of $kJ_1(k) = BJ_0(k)$.

35. $u(r,\theta) = 400 \left\{ \dfrac{r}{7a} P_1[\cos(\theta)] - \dfrac{1}{9}\left(\dfrac{r}{a}\right)^3 P_3[\cos(\theta)] - \dfrac{2}{63}\left(\dfrac{r}{a}\right)^5 P_5[\cos(\theta)] \right\}$

37. $u(r,\theta) = \tfrac{1}{2} T_0 \displaystyle\sum_{n=0}^{\infty} \{P_{n-1}[\cos(\alpha)] - P_{n+1}[\cos(\alpha)]\} \left(\dfrac{r}{a}\right)^n P_n[\cos(\theta)]$

Section 9.4

1. $$u(x,y) = \frac{64\,R}{\pi^4\,T}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{(-1)^{n+1}(-1)^{m+1}}{(2n-1)(2m-1)} \times \frac{\cos[(2n-1)\pi x/2a]\cos[(2m-1)\pi y/b]}{(2n-1)(2m-1)[(2n-1)^2/a^2+(2m-1)^2/b^2]}$$

Section 10.1

1. $1+2i$ 3. $-2/5$ 5. $2+2i\sqrt{3}$ 7. $4e^{\pi i}$ 9. $5\sqrt{2}e^{3\pi i/4}$ 11. $2e^{2\pi i/3}$

Section 10.2

1. $\pm\sqrt{2}, \quad \pm\sqrt{2}\left(\frac{1}{2}+\frac{\sqrt{3}i}{2}\right), \quad \pm\sqrt{2}\left(-\frac{1}{2}+\frac{\sqrt{3}i}{2}\right)$ 3. $i, \quad \pm\frac{\sqrt{3}}{2}-\frac{i}{2}$

5. $\pm\frac{1}{\sqrt{2}}\left(-\sqrt{\sqrt{a^2+b^2}+a}+i\sqrt{\sqrt{a^2+b^2}+a}\right)$ 7. $\pm(1+i), \quad \pm 2(1-i)$

Section 10.3

1. $u = 2-y,\ v = x$ 3. $u = x^3-3xy^2,\ v = 3x^2y-y^3$

5. $f'(z) = 3z(1+z^2)^{1/2}$ 7. $f'(z) = 2(1+4i)z-3$

9. $f'(z) = -3i(iz-1)^{-4}$ 11. $1/6$

13. $v(x,y) = 2xy + \text{constant}$ 15. $v(x,y) = x\sin(x)e^{-y} + ye^{-y}\cos(x) + \text{constant}$

Section 10.4

1. 0 3. $2i$ 5. $14/15 - i/3$

Section 10.5

1. $(e^{-2}-e^{-4})/2$ 3. $\pi/2$

Section 10.6

1. $\pi i/32$ 3. $\pi i/2$ 5. $-2\pi i$ 7. $2\pi i$ 9. 6π

Section 10.7

1. $\sum_{n=0}^{\infty}(n+1)z^n$

3. $f(z) = z^{10} - z^9 + \frac{z^8}{2} - \frac{z^7}{6} + \cdots - \frac{1}{11!z} + \cdots$
We have an essential singularity and the residue equals $-1/11!$

5. $f(z) = \frac{1}{2!} + \frac{z^2}{4!} + \frac{z^4}{6!} + \cdots$
We have a removable singularity where the value of the residue equals zero.

7. $f(z) = -\frac{2}{z} - 2 - \frac{7z}{6} - \frac{z^2}{2} - \cdots$
We have a simple pole and the residue equals -2.

9. $f(z) = \frac{1}{2}\frac{1}{z-2} - \frac{1}{4} + \frac{z-2}{8} - \cdots$
We have a simple pole and the residue equals $1/2$.

Section 10.8

1. $-3\pi i/4$ 3. $-2\pi i$. 5. $2\pi i$ 7. $2\pi i$

Section 10.11

3. $z = C\sqrt{\tau} - \pi + \pi\, i$ 5. $z = \tau^{7/4}$ or $\tau = z^{4/7}$

7. $z = \frac{a}{\pi}\cosh^{-1}(\tau), \quad 0 \le \Im[\cosh^{-1}(\tau)] \le \pi$

Section 11.3

1. $\pi e^{-|\omega/a|}/|a|$

Section 11.4

1. $-t/(1+t^2)^2$ 3. $f(t) = \frac{1}{2}e^{-t}H(t) + \frac{1}{2}e^{t}H(-t)$

5. $f(t) = e^{-t}H(t) - e^{-t/2}H(t) + \frac{1}{2}te^{-t/2}H(t)$ 7. $f(t) = ie^{-at}H(t)/2 - ie^{at}H(-t)/2$

9. $f(t) = (1 - a|t|)e^{-a|t|}/(4a)$ 11. $f(t) = (-1)^{n+1}t^{2n+1}e^{-at}H(t)/(2n+1)!$

13. $f(t) = e^{2t}H(-t) + e^{-t}H(t)$

15. $$f_+(t) = \frac{i\, e^{-at}}{R^2 - e^{2a}} + \frac{1}{2R^{t+2}} \sum_{n=-\infty}^{\infty} \frac{e^{in\pi t}}{n\pi + [\ln(R) - a]\, i}$$
$$f_-(t) = \frac{i\, e^{-at}}{R^2 e^{-2a} - 1} H(t-2) + \frac{H(t-2)}{2R^t} \sum_{n=-\infty}^{\infty} \frac{e^{in\pi t}}{n\pi + [\ln(R) - a]\, i}$$

17.
$$f(t) = -\frac{2\beta}{L} H(t) \sum_{n=1}^{\infty} (-1)^n e^{-(2n-1)\beta\gamma\pi t/2L} \times \{\gamma\cos[(2n-1)\beta\pi t/2L] + \sin[(2n-1)\beta\pi t/2L]\}$$

Section 11.6

1. $y(t) = [(t-1)e^{-t} + e^{-2t}]H(t)$ 3. $y(t) = \frac{1}{9}e^{-t}H(t) + \left[\frac{1}{9}e^{2t} - \frac{1}{3}te^{2t}\right]H(-t)$

Section 11.7

1. $u(x,y) = \frac{1}{\pi}\left[\tan^{-1}\left(\frac{1-x}{y}\right) + \tan^{-1}\left(\frac{x}{y}\right)\right]$ 3. $u(x,y) = \frac{T_0}{\pi}\left[\frac{\pi}{2} - \tan^{-1}\left(\frac{x}{y}\right)\right]$

5. $u(x,y) = \frac{T_0}{\pi}\left[\tan^{-1}\left(\frac{1-x}{y}\right) + \tan^{-1}\left(\frac{1+x}{y}\right)\right] + \frac{T_1-T_0}{2\pi}\, y \ln\left[\frac{(x-1)^2+y^2}{x^2+y^2}\right]$
$+\frac{T_1-T_0}{\pi}\, x \left[\tan^{-1}\left(\frac{1-x}{y}\right) + \tan^{-1}\left(\frac{x}{y}\right)\right]$

Section 11.8

1. $u(x,t) = \frac{1}{2}\,\mathrm{erf}\left(\frac{b-x}{\sqrt{4a^2t}}\right) + \frac{1}{2}\,\mathrm{erf}\left(\frac{b+x}{\sqrt{4a^2t}}\right)$

3. $u(x,t) = \frac{1}{2}T_0\,\mathrm{erf}\left(\frac{b-x}{\sqrt{4a^2t}}\right) + \frac{1}{2}T_0\,\mathrm{erf}\left(\frac{x}{\sqrt{4a^2t}}\right)$

Section 12.1

1. $F(s) = s/(s^2-a^2)$

3. $F(s) = 1/s + 2/s^2 + 2/s^3$

5. $F(s) = \left[1-e^{-2(s-1)}\right]/(s-1)$

7. $F(s) = 2/(s^2+1) - s/(s^2+4) + \cos(3)/s - 1/s^2$

9. $f(t) = e^{-3t}$

11. $f(t) = \frac{1}{3}\sin(3t)$

13. $f(t) = 2\sin(t) - \frac{15}{2}t^2 + 2e^{-t} - 6\cos(2t)$

17. $F(s) = 1/(2s) - sT^2/[2(s^2T^2+\pi^2)]$

Section 12.2

1. $f(t) = (t-2)H(t-2) - (t-2)H(t-3)$

3. $y'' + 3y' + 2y = H(t-1)$

5. $y'' + 4y' + 4y = tH(t-2)$

7. $y'' - 3y' + 2y = e^{-t}H(t-2)$

9. $y'' + y = \sin(t)[1 - H(t-\pi)]$

Section 12.3

1. $F(s) = 2/(s^2+2s+5)$

3. $F(s) = 2e^{-s}/s^3 + 2e^{-s}/s^2 + e^{-s}/s$

5. $F(s) = 1/(s-1)^2 + 3/(s^2-2s+10) + (s-2)/(s^2-4s+29)$

7. $F(s) = 2/(s+1)^3 + 2/(s^2-2s+5) + (s+3)/(s^2+6s+18)$

9. $F(s) = 2e^{-s}/s^3 + 2e^{-s}/s^2 + 3e^{-s}/s + e^{-2s}/s$

11. $F(s) = (1+e^{-s\pi})/(s^2+1)$

13. $F(s) = 4(s+3)/(s^2+6s+13)^2$

15. $f(t) = \frac{1}{2}t^2e^{-2t} - \frac{1}{3}t^3e^{-2t}$

17. $f(t) = e^{-t}\cos(t) + 2e^{-t}\sin(t)$

19. $f(t) = e^{-2t} - 2te^{-2t} + \cos(t)e^{-t} + \sin(t)e^{-t}$

21. $f(t) = e^{t-3}H(t-3)$

23. $f(t) = e^{-(t-1)}[\cos(t-1) - \sin(t-1)]H(t-1)$

25. $f(t) = \cos[2(t-1)]H(t-1) + \frac{1}{6}(t-3)^3e^{2(t-3)}H(t-3)$

27. $f(t) = \{\cos[2(t-1)] + \frac{1}{2}\sin[2(t-1)]\}H(t-1) + \frac{1}{6}(t-3)^3H(t-3)$

29. $f(t) = t[H(t) - H(t-a)];\quad F(s) = 1/s^2 - e^{-as}/s^2 - ae^{-as}/s$

31. $F(s) = 1/s^2 - e^{-s}/s^2 - e^{-2s}/s$

33. $F(s) = e^{-s}/s^2 - e^{-2s}/s^2 - e^{-3s}/s$

35. $Y(s) = s/(s^2+4) + 3e^{-4s}/[s(s^2+4)]$ 37. $Y(s) = e^{-(s-1)}/[(s-1)(s+1)(s+2)]$

39. $Y(s) = 5/[(s-1)(s-2)] + e^{-s}/[s^3(s-1)(s-2)]$
$+2e^{-s}/[s^2(s-1)(s-2)] + e^{-s}/[s(s-1)(s-2)]$

41. $Y(s) = 1/[s^2(s+2)(s+1)] + ae^{-as}/[(s+1)^2(s+2)]$
$-e^{-as}/[s^2(s+1)(s+2)] - e^{-as}/[s(s+1)(s+2)]$

43. $f(0) = 1$ 45. $f(0) = 0$ 47. Yes 49. No 51. No

Section 12.4

1. $F(s) = \coth\left(\frac{s\pi}{2}\right)\Big/(s^2+1)$ 3. $F(s) = \dfrac{1-(1+as)e^{-as}}{s^2(1-e^{-2as})}$

Section 12.5

1. $f(t) = e^{-t} - e^{-2t}$ 3. $f(t) = \frac{5}{4}e^{-t} - \frac{6}{5}e^{-2t} - \frac{1}{20}e^{3t}$

5. $f(t) = e^{-2t}\cos\left(t + \frac{3\pi}{2}\right)$ 7. $f(t) = 2.3584\cos(4t + 0.5586)$

9. $f(t) = \frac{1}{2} + \frac{\sqrt{2}}{2}\cos\left(2t + \frac{5\pi}{4}\right)$

Section 12.6

11. $f(t) = e^t - t - 1$

Section 12.7

1. $f(t) = 1 + 2t$ 3. $f(t) = t + t^2/2$

5. $f(t) = t^3 + t^5/20$ 7. $f(t) = t^2 - t^4/3$

9. $f(t) = 5e^{2t} - 4e^t - 2te^t$ 11. $f(t) = (1-t)^2e^{-t}$

13. $f(t) = a\left[1 - e^{\pi t}\,\mathrm{erfc}(\sqrt{\pi t})\right]/2$ 15. $f(t) = \frac{1}{2}t^2$

17. $x(t) = \left\{e^{c^2t}\left[1 + \mathrm{erf}\left(c\sqrt{t}\,\right)\right] - c^2t - 1 - 2c\sqrt{t/\pi}\right\}/c^2$

19. $f(t) = a + a^2t + \frac{1}{2}a\,\mathrm{erf}(\sqrt{at}\,) + a^3t\,\mathrm{erf}(\sqrt{at}\,) + a^2\sqrt{t}e^{-a^2t}/\sqrt{\pi}$

Section 12.8

1. $y(t) = \frac{5}{4}e^{2t} - \frac{1}{4} + \frac{1}{2}t$ 3. $y(t) = e^{3t} - e^{2t}$

5. $y(t) = -\frac{3}{4}e^{-3t} + \frac{7}{4}e^{-t} + \frac{1}{2}te^{-t}$ 7. $y(t) = \frac{3}{4}e^{-t} + \frac{1}{8}e^t - \frac{7}{8}e^{-3t}$

9. $y(t) = (t-1)H(t-1)$ 11. $y(t) = e^{2t} - e^t + \left[\frac{1}{2} + \frac{1}{2}e^{2(t-1)} - e^{t-1}\right]H(t-1)$

13. $y(t) = \left[1 - e^{-2(t-2)} - 2(t-2)e^{-2(t-2)}\right]H(t-2)$

15. $y(t) = \left[\frac{1}{3}e^{2(t-2)} - \frac{1}{2}e^{t-2} + \frac{1}{6}e^{-(t-2)}\right]H(t-2)$

17. $y(t) = 1 - \cos(t) - [1 - \cos(t-T)]\,H(t-T)$

19. $y(t) = e^{-t} - \frac{1}{4}e^{-2t} - \frac{3}{4} + \frac{1}{2}t - \left[e^{-(t-a)} - \frac{1}{4}e^{-2(t-a)} - \frac{3}{4} + \frac{1}{2}(t-a)\right]H(t-a)$
$+a\left[\frac{1}{2}e^{-2(t-a)} + (t-a)e^{-(t-a)} - \frac{1}{2}\right]H(t-a)$

21. $y(t) = te^t + 3(t-2)e^{t-2}H(t-2)$

23. $y(t) = 3\left[e^{-2(t-2)} - e^{-3(t-2)}\right]H(t-2) + 4\left[e^{-3(t-5)} - e^{-2(t-5)}\right]H(t-5)$

25. $x(t) = \cos(\sqrt{2}t)e^{3t} - \frac{1}{\sqrt{2}}\sin(\sqrt{2}t)e^{3t}$; $\quad y(t) = \frac{3}{\sqrt{2}}\sin(\sqrt{2}t)e^{3t}$

27. $x(t) = t - 1 + e^{-t}\cos(t)$, $\qquad y(t) = t^2 - t + e^{-t}\sin(t)$

29. $x(t) = 3F_1 - 2F_2 - F_1\cosh(t) + F_2e^t - 2F_1\cos(t) + F_2\cos(t) - F_2\sin(t)$
$y(t) = F_2 - 2F_1 + F_1e^{-t} - F_2\cos(t) + F_1\cos(t) + F_1\sin(t)$

Section 12.9

1. $f(t) = (2-t)e^{-2t} - 2e^{-3t}$

3. $f(t) = \left(\frac{1}{4}t^2 - \frac{1}{4}t + \frac{1}{8}\right)e^{2t} - \frac{1}{8}$

5. $f(t) = \left[\frac{1}{2}(t-1) - \frac{1}{4} + \frac{1}{4}e^{-2(t-1)}\right]H(t-1)$

7. $$f(t) = \frac{e^{-bt}}{\cosh(ab)} - 8ab\sum_{n=1}^{\infty}(-1)^n\frac{\sin[(2n-1)\pi t/(2a)]}{4a^2b^2 + (2n-1)^2\pi^2} + 4\sum_{n=1}^{\infty}(-1)^n\frac{(2n-1)\pi\cos[(2n-1)\pi t/(2a)]}{4a^2b^2 + (2n-1)^2\pi^2}$$

Section 12.10

1. $$u(x,t) = \frac{4}{\pi^2}\sum_{m=1}^{\infty}\frac{\sin[(2m-1)\pi x]\sin[(2m-1)\pi t]}{(2m-1)^2}$$

3. $u(x,t) = \sin(\pi x)\cos(\pi t) - \sin(\pi x)\sin(\pi t)/\pi$

5. $$u(x,t) = c\sum_{n=0}^{\infty}f(t - x/c - 2nL/c)H(t - x/c - 2nL/c) + c\sum_{m=1}^{\infty}f(t + x/c - 2mL/c)H(t + x/c - 2mL/c)$$

7. $u(x,t) = xt - te^{-x} + \sinh(t)e^{-x} + \left[1 - e^{-(t-x)} + t - x - \sinh(t-x)\right]H(t-x)$

9. $$u(x,t) = \frac{gx}{\omega^2} - \frac{2g\omega^2}{L}\sum_{n=1}^{\infty}\frac{\sin(\lambda_n x)\cos(\lambda_n t)}{\lambda_n^2(\omega^4 + \omega^2/L + \lambda_n^2)\sin(\lambda_n L)},$$ where λ_n is the nth root of $\lambda = \omega^2\cot(\lambda L)$.

11. $$u(x,t) = E - \frac{4E}{\pi}\sum_{n=1}^{\infty}\frac{1}{2n-1}\sin\left[\frac{(2n-1)\pi x}{2\ell}\right]\cos\left[\frac{(2n-1)c\pi t}{2\ell}\right],$$ or
$$u(x,t) = E\sum_{n=0}^{\infty}(-1)^n H\left(t - \frac{x + 2n\ell}{c}\right) + E\sum_{n=0}^{\infty}(-1)^n H\left\{t - \frac{[(2n+2)\ell - x]}{c}\right\}.$$

13. $$p(x,t) = p_0 - \frac{4\rho u_0 c}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{2n-1}\sin\left[\frac{(2n-1)\pi x}{2L}\right]\sin\left[\frac{(2n-1)c\pi t}{2L}\right]$$

15. $$u(r,t) = \frac{ap_0}{\rho r[(\beta/\sqrt{2} - \alpha)^2 + \beta^2]}\left\{e^{-\beta\tau/\sqrt{2}}\left[\left(\frac{1}{\sqrt{2}} - \frac{\alpha}{\beta}\right)\sin(\beta\tau) + \cos(\beta\tau)\right]\right.$$

$$-e^{-\alpha\tau}\Big\}H(\tau), \text{ where } \tau = t - (r-a)/c.$$

17. $u(x,t) = \dfrac{2t}{3} + \dfrac{x^2}{2} - \dfrac{1}{6} - 2\sum_{n=1}^{\infty}(-1)^n \cos(n\pi x)\left[\dfrac{\cos(n\pi t)}{n^2\pi^2} + \dfrac{2\sin(n\pi t)}{n^3\pi^3}\right]$

Section 12.11

1. $u(x,t) = T_0\left(1 - e^{-a^2 t}\right)$

3. $u(x,t) = x + \dfrac{2}{\pi}\sum_{n=1}^{\infty}\dfrac{(-1)^n}{n}\sin(n\pi x)e^{-n^2\pi^2 t}$

5. $u(x,t) = \dfrac{x(1-x)}{2} - \dfrac{4}{\pi^3}\sum_{m=1}^{\infty}\dfrac{\sin[(2m-1)\pi x]}{(2m-1)^3}e^{-(2m-1)^2\pi^2 t}$

7. $u(x,t) = x\,\text{erfc}\left(\dfrac{x}{2\sqrt{t}}\right) - 2\sqrt{\dfrac{t}{\pi}}\exp\left(-\dfrac{x^2}{4t}\right)$

9. $u(x,t) = \dfrac{u_0}{2}e^{-\delta x}\text{erfc}\left(\dfrac{x}{2a\sqrt{t}} + \dfrac{a(1-\delta)\sqrt{t}}{2}\right) + \dfrac{u_0}{2}e^{-x}\text{erfc}\left(\dfrac{x}{2a\sqrt{t}} - \dfrac{a(1-\delta)\sqrt{t}}{2}\right)$

11. $u(x,t) = \frac{1}{2}e^{\sigma^2 t - \beta t}\left[\dfrac{e^{-\sigma x}}{\rho+\sigma}\text{erfc}\left(\dfrac{x}{2\sqrt{t}} - \sigma\sqrt{t}\right) + \dfrac{e^{\sigma x}}{\rho-\sigma}\text{erfc}\left(\dfrac{x}{2\sqrt{t}} + \sigma\sqrt{t}\right)\right]$
$\qquad - \dfrac{\rho}{\rho^2-\sigma^2}e^{\rho x + \rho^2 t - \beta t}\text{erfc}\left(\dfrac{x}{2\sqrt{t}} + \rho\sqrt{t}\right)$

13. $u(x,t) = \dfrac{t(L-x)}{L} + \dfrac{Px(x-L)}{2a^2} - \dfrac{x(x-L)(x-2L)}{6a^2L}$
$\qquad - \dfrac{2PL^2}{a^2\pi^3}\sum_{n=1}^{\infty}\dfrac{(-1)^n}{n^3}\sin\left(\dfrac{n\pi x}{L}\right)\exp\left(-\dfrac{a^2n^2\pi^2 t}{L^2}\right)$
$\qquad + \dfrac{2(P+1)L^2}{a^2\pi^3}\sum_{n=1}^{\infty}\dfrac{1}{n^3}\sin\left(\dfrac{n\pi x}{L}\right)\exp\left(-\dfrac{a^2n^2\pi^2 t}{L^2}\right)$

15. $u(x,t) = \dfrac{4q}{\pi}\sum_{n=1}^{\infty}\dfrac{(-1)^n\cos[(2n-1)\pi x/2L]}{(2n-1)[\alpha q - (2n-1)^2\pi^2a^2/4L^2]}$
$\qquad \times\left\{1 - \exp[\alpha q t - (2n-1)^2\pi^2a^2t/4L^2]\right\}$

17. $u(r,t) = u_0\left[1 - \dfrac{b-\beta}{r}f(r,t)\right] + \displaystyle\int_0^t\left[1 - \dfrac{b-\beta}{r}f(r,t-\tau)\right]q(\tau)\,d\tau,$
where $f(r,t) = \text{erfc}\left(\dfrac{r-b}{2a\sqrt{t}}\right) - \exp\left(\dfrac{r-b}{\beta} + \dfrac{a^2t}{\beta^2}\right)\text{erfc}\left(\dfrac{a\sqrt{t}}{\beta} + \dfrac{r-b}{2a\sqrt{t}}\right)$.

19. $y(t) = \dfrac{4\mu A\omega^2}{mL}\sum_{n=1}^{\infty}\dfrac{\lambda_n e^{\lambda_n t}}{\lambda_n^4 - (\frac{2\mu}{mL})(1+\frac{2\mu L}{m\nu})\lambda_n^3 + 2\omega^2\lambda_n^2 + \frac{6\omega^2\mu}{mL}\lambda_n + \omega^4},$
where λ_n is the nth root of $\lambda^2 + 2\mu\lambda^{3/2}\coth\left(L\sqrt{\lambda/\nu}\right)/(m\sqrt{\nu}) + \omega^2 = 0$.

21. $u(x,t) = \dfrac{x}{a+1} + 2\sum_{n=1}^{\infty}\dfrac{\sin(\lambda_n x)\exp(-\lambda_n^2 t)}{[3a+3+\lambda_n^2/(3a)]\sin(\lambda_n)},$
where λ_n is the nth root of $\lambda\cot(\lambda) = (3a+\lambda^2)/3a$.

23. $u(x,t) = 1 - e^{-bt} \int_0^{ax} e^{-\eta} I_0\left(2\sqrt{bt\eta}\right)\, d\eta$

25. $u(r,t) = \dfrac{a^2 - r^2}{4} - 2a^2 \sum_{n=1}^{\infty} \dfrac{J_0(k_n r/a)}{k_n^3\, J_1(k_n)} e^{-k_n^2 t/a^2}$,
where k_n is the nth root of $J_0(k) = 0$.

27. $u(r,t) = k \left[t - \dfrac{b^2 - r^2}{4a^2} + \dfrac{2}{a^2 b} \sum_{n=1}^{\infty} \dfrac{J_0(\kappa_n r)}{\kappa_n^3 J_1(\kappa_n b)} \right]$,
where k_n is the nth root of $J_0(\kappa b) = 0$.

Section 12.13

1. $u(x,y) = \dfrac{4}{\pi} \sum_{m=1}^{\infty} \dfrac{1}{2m-1} \exp\left[-\dfrac{(2m-1)\pi x}{a} \right] \sin\left[\dfrac{(2m-1)\pi y}{a} \right]$

Section 13.1

1. $F(z) = 2z/(2z-1)$ if $|z| > 1/2$

3. $F(z) = (z^6 - 1)/(z^6 - z^5)$ if $|z| > 0$

5. $F(z) = (a^2 + a - z)/[z(z-a)]$ if $|z| > a$.

Section 13.2

1. $F(z) = zTe^{aT}/(ze^{aT} - 1)^2$

3. $F(z) = z(z+a)/(z-a)^3$

5. $F(z) = [z - \cos(1)]/\{z[z^2 - 2z\cos(1) + 1]\}$

7. $F(z) = z[z\sin(\theta) + \sin(\omega_0 T - \theta)]/[z^2 - 2z\cos(\omega_0 T) + 1]$

9. $F(z) = z/(z+1)$

11. $f_n * g_n = n + 1$

13. $f_n * g_n = 2^n/n!$

Section 13.3

1. $f_0 = 0.007143,\ f_1 = 0.08503,\ f_2 = 0.1626,\ f_3 = 0.2328$

3. $f_0 = 0.09836,\ f_1 = 0.3345,\ f_2 = 0.6099,\ f_3 = 0.7935$

5. $f_n = 8 - 8\left(\frac{1}{2}\right)^n - 6n\left(\frac{1}{2}\right)^n$

7. $f_n = (1 - \alpha^{n+1})/(1 - \alpha)$

9. $f_n = \left(\frac{1}{2}\right)^{n-10} H_{n-10} + \left(\frac{1}{2}\right)^{n-11} H_{n-11}$

11. $f_n = \frac{1}{9}(6n-4)(-1)^n + \frac{4}{9}\left(\frac{1}{2}\right)^n$

13. $f_n = a^n/n!$

Section 13.4

1. $y_n = 1 + \frac{1}{6}n(n-1)(2n-1)$

3. $y_n = \frac{1}{2}n(n-1)$

5. $y_n = \frac{1}{6}\left[5^n - (-1)^n\right]$

7. $y_n = (2n-1)\left(\frac{1}{2}\right)^n + \left(-\frac{1}{2}\right)^n$

9. $y_n = 2^n - n - 1$

11. $x_n = 2 + (-1)^n; y_n = 1 + (-1)^n$

13. $x_n = 1 - 2(-6)^n; y_n = -7(-6)^n$

Section 13.5

1. marginally stable

3. unstable

Section 14.1

7. $\hat{x}(t) = \dfrac{1}{\pi}\ln\left|\dfrac{t+a}{t-a}\right|$

Section 14.2

5. $w(t) = u(t) * v(t) = \pi e^{-1}\sin(t)$

Section 14.3

1. $z(t) = e^{i\omega t}$

Section 14.4

3. $x(t) = \dfrac{1-t^2}{(1+t^2)^2};\quad \hat{x}(t) = \dfrac{2t}{(1+t^2)^2}$

Section 15.2

1. $G(s) = 1/(s+k)$, $g(t|0) = e^{-kt}$
   $g(t|\tau) = e^{-k(t-\tau)}H(t-\tau)$, $a(t) = \left(1-e^{-kt}\right)/k$

3. $G(s) = 1/(s^2+4s+3)$, $g(t|0) = \frac{1}{2}(e^{-t}-e^{-3t})$
   $g(t|\tau) = \frac{1}{2}\left[e^{-(t-\tau)}-e^{-3(t-\tau)}\right]H(t-\tau)$
   $a(t) = \frac{1}{6}e^{-3t} - \frac{1}{2}e^{-t} + \frac{1}{3}$

5. $G(s) = 1/[(s-2)(s-1)]$, $g(t|0) = e^{2t} - e^{t}$
   $g(t|\tau) = \left[e^{2(t-\tau)} - e^{t-\tau}\right]H(t-\tau)$, $a(t) = \frac{1}{2} + \frac{1}{2}e^{2t} - e^{t}$

7. $G(s) = 1/(s-9)^2$, $g(t|0) = \frac{1}{3}\sinh(3t)$
   $g(t|\tau) = \frac{1}{3}\sinh[3(t-\tau)]H(t-\tau)$, $a(t) = \frac{1}{9}\left[\cosh(3t)-1\right]$

9. $G(s) = 1/[s(s-1)]$, $g(t|0) = e^{t} - 1$
   $g(t|\tau) = \left[e^{t-\tau}-1\right]H(t-\tau)$, $a(t) = e^{t} - t - 1$

11.
$$g(x|\xi) = \frac{(1+x_<)(L-1-x_>)}{L},$$
and
$$g(x|\xi) = -\frac{2e^{x+\xi}}{e^{2L}-1} + \frac{2L^3}{\pi^2}\sum_{n=1}^{\infty}\frac{\varphi_n(\xi)\varphi_n(x)}{n^2(n^2\pi^2+L^2)},$$
where $\varphi_n(x) = \sin(n\pi x/L) + n\pi\cos(n\pi x/L)/L$.

13.
$$g(x|\xi) = \frac{\sinh(kx_<)\sinh[k(L-x_>)]}{k\,\sinh(kL)},$$

and

$$g(x|\xi) = 2L \sum_{n=1}^{\infty} \frac{\sin(n\pi\xi/L)\sin(n\pi x/L)}{n^2\pi^2 + k^2L^2}.$$

15.

$$g(x|\xi) = \frac{\sinh(kx_<)\{k\cosh[k(x_> - L)] - \sinh[k(x_> - L)]\}}{k\sinh(kL) + k^2\cosh(kL)},$$

and

$$g(x|\xi) = 2\sum_{n=1}^{\infty} \frac{(1+k_n^2)\sin(k_n\xi)\sin(k_n x)}{[1+(1+k_n^2)L](k_n^2+k^2)},$$

where k_n is the nth root of $\tan(kL) = -k$.

17.

$$g(x|\xi) = \frac{[a\sinh(kx_<) - k\cosh(kx_<)]\cosh[k(L - x_>)]}{k[a\cosh(kL) - k\sinh(kL)]},$$

and

$$g(x|\xi) = 2\sum_{n=1}^{\infty} \frac{(a^2+k_n^2)\cos[k_n(\xi - L)]\cos[k_n(x-L)]}{[(a^2+k_n^2)L - a](k_n^2+k^2)},$$

where k_n is the nth root of $k\,\tan(kL) = -a$.

Section 15.4

3.

$$g(x,t|\xi,\tau) = \frac{t-\tau}{L}H(t-\tau) + \frac{2}{\pi}H(t-\tau)\sum_{n=1}^{\infty}\frac{1}{n}\cos\left(\frac{n\pi\xi}{L}\right)\cos\left(\frac{n\pi x}{L}\right)\sin\left[\frac{n\pi(t-\tau)}{L}\right]$$

5.

$$u(x,t) = 2\sum_{n=1}^{\infty}\sin\left(\frac{n\pi x}{L}\right)\left\{\frac{n\pi}{L^2+n^2\pi^2}\left[e^{-t} - \cos\left(\frac{n\pi t}{L}\right)\right] + \frac{L}{L^2+n^2\pi^2}\sin\left(\frac{n\pi t}{L}\right)\right\}$$
$$+ 2\sin\left(\frac{\pi x}{L}\right)\cos\left(\frac{\pi t}{L}\right)$$
$$+ \frac{4L}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^2}\sin\left[\frac{(2m-1)\pi x}{L}\right]\sin\left[\frac{(2m-1)\pi t}{L}\right]$$

7.

$$u(x,t) = 1 - \frac{t^2}{2L} - \frac{2L}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\cos\left(\frac{n\pi x}{L}\right)\left[1 - \cos\left(\frac{n\pi t}{L}\right)\right]$$

Section 15.5

3.

$$g(x,t|\xi,\tau) = \frac{2}{L}\left\{\sum_{n=1}^{\infty}\sin\left[\frac{(2n-1)\pi\xi}{2L}\right]\sin\left[\frac{(2n-1)\pi x}{2L}\right]\exp\left[-\frac{(2n-1)^2\pi^2(t-\tau)}{4L^2}\right]\right\}$$
$$\times H(t-\tau)$$

5.

$$u(x,t) = 2\pi \sum_{n=1}^{\infty} \frac{n}{n^2\pi^2 - L^2} \sin\left(\frac{n\pi x}{L}\right) \left[e^{-t} - \exp\left(-\frac{n^2\pi^2 t}{L^2}\right)\right]$$
$$+ \frac{4}{\pi} \sum_{m=1}^{\infty} \frac{1}{2m-1} \sin\left[\frac{(2m-1)\pi x}{L}\right] \exp\left[-\frac{(2m-1)^2\pi^2 t}{L^2}\right]$$

7.

$$u(x,t) = 1 - \frac{t}{L} - \frac{2L}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \cos\left(\frac{n\pi x}{L}\right) \left[1 - \exp\left(-\frac{n^2\pi^2 t}{L^2}\right)\right]$$

Section 15.6

1.

$$g(x,y|\xi,\eta) = \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \exp\left(-\frac{n\pi}{a}|y-\eta|\right) \sin\left(\frac{n\pi\xi}{a}\right) \sin\left(\frac{n\pi x}{a}\right)$$

5.

$$g(r,\theta|\rho,\theta') = \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} r_<^{n\pi/\beta} r_>^{-n\pi/\beta} \sin\left(\frac{n\pi\theta'}{\beta}\right) \sin\left(\frac{n\pi\theta}{\beta}\right)$$

7.

$$g(r,z|\rho,\zeta) = \frac{2}{\pi a^2 L} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{J_0(k_m\rho/a)J_0(k_m r/a)}{\pi a^2 L J_1^2(k_m)(k_m^2/a^2 + n^2\pi^2/L^2)} \sin\left(\frac{n\pi\zeta}{L}\right) \sin\left(\frac{n\pi z}{L}\right)$$

Section 16.2

1. (a) $S = \{HH, HT, TH, TT\}$ (b) $S = \{ab, ac, ba, bc, ca, cb\}$

(c) $S = \{aa, ab, ac, ba, bb, bc, ca, cb, cc\}$ (d) $S = \{bbb, bbg, bgb, bgg, ggb, ggg, gbb, gbg\}$

(e) $S = \{bbb, bbg, bgb, bgg, ggb, ggg, gbb, gbg\}$

3. 1/3 5. 1/3 7. 2/13 9. 1/720, 1/120

11. 1/2 13. 1/2 15. 9/16

Section 16.3

1. $F_X(x) = \begin{cases} 0, & x < 0, \\ 1-p, & 0 \le x < 1, \\ 1, & 1 \le x. \end{cases}$ 3. 27

Section 16.4

1. $F_X(x) = \begin{cases} 0, & x \le 0, \\ 1 - e^{-\lambda x}, & 0 < x. \end{cases}$ 3. $F_X(x) = \begin{cases} 0, & x < -1, \\ (1+x)^2/2, & -1 \le x < 0, \\ 1 - (x-1)^2/2, & 0 \le x < 1, \\ 1, & 1 \le x. \end{cases}$

Section 16.5

1. $E(X) = \frac{1}{2}$, and $\text{Var}(X) = \frac{1}{4}$ 3. $k = 3/4$, $E(X) = 1$, and $\text{Var}(X) = \frac{1}{5}$

5. $\phi_X(\omega) = \left(pe^{i\omega} + q\right)^n$, $\mu_X = np$, $\text{Var}(X) = npq$

7. $\phi_X(\omega) = p/(1 - qe^{\omega i})$, $\mu_X = q/p$, $\text{Var}(X) = q/p^2$

Section 16.6

1. (a) 1/16, (b) 1/4, (c) 15/16, (d) 1/16 5. $P(X > 0) = 0.01$, and $P(X > 1) = 9 \times 10^{-5}$

7. $P(T < 150) = \frac{1}{3}$, and $P(X = 3) = 0.1646$

Section 16.7

1.

$$p_{XY}[x_i, y_j] = \frac{\binom{7}{x_i}\binom{8}{y_j}\binom{5}{5 - x_i - y_j}}{\binom{20}{5}},$$

where $x_i = 0, 1, 2, 3, 4, 5$, $y_j = 0, 1, 2, 3, 4, 5$ and $0 \le x_i + y_j \le 5$.

Section 17.1

1. $\mu_X(t) = 0$, and $\sigma_X^2(t) = \cos(\omega t)$

3. For $t_1 = t_2$, $R_X(t_1, t_2) = p$; for $t_1 \ne t_2$, $R_X(t_1, t_2) = p^2$. For $t_1 = t_2$, $C_X(t_1, t_2) = p(1-p)$; for $t_1 \ne t_2$, $C_X(t_1, t_2) = 0$.

Section 17.4

1.

$$P^n = \begin{pmatrix} 2/3 + (1/3)(1/4)^n & 1/3 - (1/3)(1/4)^n \\ 2/3 - (2/3)(1/4)^n & 1/3 + (2/3)(1/4)^n \end{pmatrix}. \qquad P^\infty = \begin{pmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{pmatrix}.$$

Section 18.3

1. $E(X) = 0, \qquad \text{Var}(X) = E(X^2) = (b^3 - a^3)/3$

3. $E(X) = 0, \qquad \text{Var}(X) = E(X^2) = (b^2 - a^2)/2$

Section 18.5

1. $X(t) = e^{t/2}B(t) + X_0$ 3. $X(t) = B^2(t) + tB(t) + X_0$ 5. $X(t) = tB^2(t) - t^2/2 + X_0$

7. $e^{t/(RC)}Q(t) - Q(0) = \frac{1}{R}\int_0^t e^{\eta/(RC)}V(\eta)\,d\eta + \frac{1}{R}\int_0^t e^{\eta/(RC)}\alpha(\eta)\,d\eta$

9. $X(t) = X(0)e^{4t} + \frac{1}{4}\left(1 - e^{-4t}\right) + 2\int_0^t e^{4(t-\eta)}\, dB(\eta)$

11. $X(t) = X_0 e^t + e^t - 1 + \frac{1}{2}e^t[B^2(t) - t]$

13. $X(t) = e^{t/2}X_0 - 2t - 4 + 5e^{t/2} - e^t \cos[B(t)]$

17.
$$X(t) = X_0 \exp\left[\tfrac{1}{8}\sin(2t) + \sin(t) - \tfrac{1}{4}t + \int_0^t \sin(\eta)\, dB(\eta)\right]$$

19.
$$X(t) = X_0 \exp\left[(t+1)\ln(t+1) - t - \tfrac{1}{6}t^3 + \int_0^t \eta\, dB(\eta)\right]$$

Index