Semester – VI

		L T P C
16MMU601A	**MATHEMATICAL MODELLING**	**6 2 0 6**

Scope: On successful completion of course the learners gain about the differential equations, Queuing models and its applications.

Objectives: To enable the students to learn and gain knowledge about Bessel's equation, Legendre's equation , Simulation Modeling and applications to Traffic Flow.

UNIT I

Power series solution of a differential equation about an ordinary point, solution about a regular singular point, Bessel's equation and Legendre's equation, Laplace transform and inverse transform, application to initial value problem up to second order.

UNIT II

Monte Carlo Simulation Modeling: simulating deterministic behavior (area under a curve, volume under a surface), Generating Random Numbers: middle square method, linear congruence.

UNIT III

Queuing Models: harbor system, morning rush hour, Overview of optimization modeling, Linear Programming Model: geometric solution algebraic solution, simplex method, sensitivity analysis.

UNIT IV

Applications of differential equations: the vibrations of a mass on a spring, mixture problem, free damped motion, forced motion, resonance phenomena, electric circuit problem, mechanics of simultaneous differential equations.

UNIT V

Applications to Traffic Flow. Vibrating string, vibrating membrane, conduction of heat in solids, gravitational potential, conservation laws.

SUGGESTED READINGS

TEXT BOOKS

1. Shepley L. Ross, (1984). Differential Equations, Fourth Edition, John Wiley and Sons , New York**.(For Unit-I,II & III)**

2. Sneddon I., (2006). Elements of Partial Differential Equations, McGraw-Hill, International Edition, New York. **.(For Unit-IV & V)**

REFERENCES

1. Tyn Myint-U and Lokenath Debnath, (2006). Linear Partial Differential Equation for Scientists and Engineers, Springer.

2. Frank R. Giordano, Maurice D. Weir and William P. Fox, (2003). A First Course in Mathematical Modeling, Thomson Learning, London and New York.

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)achi Main Road, Eachanari (Po), Coimbatore –641 021.
Department of Mathematics

LECTURE PLAN

Subject: Mathematical Modeling **Subject Code: 16MMU601A**

S. No.	Lecture Duration	Topic to be covered	Support Material
		Unit – I	
1	1	Power series solution of a differential equation about an ordinary point	T1: chap-6 Pg.No:221 - 229
2	1	Continuation on Power series solution of a differential equation about an ordinary point problems	T1: chap-6 Pg.No:230 - 232
3	1	Tutorial	
4	1	Solution about a regular singular point	T1: chap-6 Pg.No:233 - 239
5	1	Continuation on Solution about a regular singular point problems	T1: chap-6 Pg.No:240 - 249
6	1	Tutorial	
7	1	Bessel's equation Derivation	T1: chap-6 Pg.No:252 - 254
8	1	Continuation on Bessel's equation problems	T1: chap-6 Pg.No:255 - 257
9	1	Tutorial	
10	1	Legendre's equation Derivation	T1: chap-6 Pg.No:258 - 260
11	1	Continuation on Legendre's equation problems	T1: chap-6 Pg.No:261 - 263
12	1	Tutorial	
13	1	Laplace transform and inverse transform	T1: chap-7 Pg.No:411-412
14	1	Continuation on Laplace transform and inverse transform problems	T1: chap-7 Pg.No:441-449
15	1	Application to initial value problem up to second order	
16	1	Recapitulation and discussion of possible questions on unit I	
Total No. of Lecture hours planned – 16 Hours			
T1: Shepley L. Ross, (1984). Differential Equations, Fourth Edition, John Wiley and Sons , New York			
		Unit – II	
1	1	Monte Carlo Simulation Modeling	R2: chap-5 Pg.No:187
2	1	simulating deterministic behavior area under a curve	R2: chap-5 Pg.No:188
3	1	Tutorial	
4	1	Continuation on simulating deterministic behavior area under a curve examples	R2: chap-5 Pg.No:188
5	1	simulating deterministic behavior Volume under a surface	R2: chap-5 Pg.No:189
6	1	Tutorial	
7	1	Continuation on simulating deterministic behavior Continuation volume under a surface examples	R2: chap-5 Pg.No:1890
8	1	Generating Random Numbers	R2: chap-5 Pg.No:191
9	1	Tutorial	
10	1	Middle square method	R2: chap-5 Pg.No:192
11	1	Continuation on middle square method examples	R2: chap-5 Pg.No:192
12	1	Tutorial	
13	1	Linear congruence	R2: chap-5 Pg.No:193

S. No.	Lecture Duration	Topic to be covered	Support Material
14	1	Continuation on Linear congruence	R2: chap-5 Pg.No:194-195
15	1	Continuation on Linear congruence	R2: chap-5 Pg.No:196-197
16	1	Recapitulation and discussion of possible questions on unit II	
Total No. of Lecture hours planned – 16 Hours			
R2: Frank R. Giordano, Maurice D. Weir and William P. Fox, (2003). A First Course in Mathematical Modeling,Thomson Learning, London and, New York			
Unit – III			
1	1	Queuing Models	R2: chap-5 Pg.No:213
2	1	Harbor system	R2: chap-5 Pg.No:214 - 215
3	1	Continuation on Harbor system	R2: chap-5 Pg.No:216 - 219
4	1	Tutorial	
5	1	morning rush hour	R2: chap-5 Pg.No:219 - 220
6	1	Continuation on morning rush hour	R2: chap-5 Pg.No:221
7	1	Tutorial	
8	1	Overview of optimization modeling	R2: chap-7 Pg.No:255
9	1	Linear Programming Model Geometric solution	R2: chap-7 Pg.No:256 - 262
10	1	Tutorial	
11	1	Linear Programming algebraic solution	R2: chap-7 Pg.No:265 - 269
12	1	simplex method	R2: chap-7 Pg.No:269 - 272
13	1	Tutorial	
14	1	Continuation on simplex method	R2: chap-7 Pg.No:273 - 278
15	1	sensitivity analysis	R2: chap-7 Pg.No:279 - 284
16	1	Recapitulation and discussion of possible questions on unit III	
Total No. of Lecture hours planned – 16 Hours			
R2: Frank R. Giordano, Maurice D. Weir and William P. Fox, (2003). A First Course in Mathematical Modeling,Thomson Learning, London and, New York			
Unit – IV			
1	1	Applications of differential equations	T1: chap-5 Pg.No:179
2	1	The vibrations of a mass on a spring	T1: chap-5 Pg.No:179 - 180
3	1	Tutorial	
4	1	mixture problem	T1: chap-5 Pg.No:181 - 182
5	1	force damped motion	T1: chap-5 Pg.No:199 - 202
6	1	Tutorial	
7	1	forced motion	T1: chap-5 Pg.No: 202 - 205
8	1	resonance phenomena	T1: chap-5 Pg.No: 206 - 208
9	1	Tutorial	
10	1	Continuation on resonance phenomena	T1: chap-5 Pg.No: 209 - 210
11	1	Electric circuit problem	T1: chap-5 Pg.No: 211 - 212
12	1	Tutorial	
13	1	Continuation on electric circuit problem	T1: chap-5 Pg.No: 213 - 215

S. No.	Lecture Duration	Topic to be covered	Support Material
14	1	mechanics of simultaneous differential equations.	T1: chap-5 Pg.No: 216 - 217
15	1	Continuation on mechanics of simultaneous differential equations.	T1: chap-5 Pg.No: 218 - 220
16	1	Recapitulation and discussion of possible questions on unit IV	
Total No. of Lecture hours planned – 16 Hours			
T1:Shepley L. Ross, (1984). Differential Equations, Fourth Edition, John Wiley and Sons , New York			
Unit – V			
1	1	Applications to Traffic Flow	R1: chap-3 Pg.No: 65
2	1	Vibrating string	R1: chap-3 Pg.No: 65
3	1	Tutorial	
4	1	vibrating membrane	R1: chap-3 Pg.No: 67 - 69
5	1	conduction of heat in solids	R1: chap-3 Pg.No: 70 - 72
6	1	Tutorial	
7	1	Continuation on heat in solids	R1: chap-3 Pg.No: 73 - 75
8	1	gravitational potential	R1: chap-3 Pg.No: 76 - 77
9	1	Tutorial	
10	1	Continuation on gravitational potential	R1: chap-3 Pg.No: 77 - 78
11	1	conservation laws	R1: chap-3 Pg.No: 79 - 80
12	1	Tutorial	
13	1	Recapitulation and discussion of possible questions on unit V	
14	1	Discussion of previous ESE question papers	
15	1	Discussion of previous ESE question papers	
16	1	Discussion of previous ESE question papers	
Total No. of Lecture hours planned – 16 Hours			
R1: Tyn Myint and Lokenath Debnath,(2006). Linear Partial Differential Equation for Scientists and Engineers, Springer.			

SUGGESTED READINGS

TEXT BOOKS

1. Shepley L. Ross, (1984). Differential Equations, Fourth Edition, John Wiley and Sons , New York.(For Unit-I,II & III)

2. Sneddon I., (2006). Elements of Partial Differential Equations, McGraw-Hill, International Edition,New York. .(For Unit-IV & V)

REFERENCES

1. Tyn Myint-U and Lokenath Debnath, (2006). Linear Partial Differential Equation for Scientists and Engineers, Springer.

2. Frank R. Giordano, Maurice D. Weir and William P. Fox, (2003). A First Course in Mathematical Modeling, Thomson Learning, London and New York.

Name and Signature
e Student Representa

Name and Signature
of Course Faculty

S. No.	Lecture Duration	Topic to be covered	Support Material

and Signature of Clas

Name and Signature of Coordinato

Head of the Department

r

6.1 POWER SERIES SOLUTIONS ABOUT AN ORDINARY POINT

A. *Basic Concepts and Results*

Consider the second-order homogeneous linear differential equation

$$a_0(x)\frac{d^2y}{dx^2} + a_1(x)\frac{dy}{dx} + a_2(x)y = 0, \tag{6.1}$$

and suppose that this equation has no solution that is expressible as a finite linear combination of known elementary functions. Let us assume, however, that it does have a solution that can be expressed in the form of an infinite series. Specifically, we assume that it has a solution expressible in the form

$$c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + \cdots = \sum_{n=0}^{\infty} c_n(x - x_0)^n, \tag{6.2}$$

where $c_0, c_1, c_2, \ldots$ are constants. An expression of the form (6.2) is called a *power series* in $x - x_0$. We have thus assumed that the differential equation (6.1) has a so-called *power series solution* of the form (6.2). Assuming that this assumption is valid, we can proceed to determine the coefficients $c_0, c_1, c_2, \ldots$ in (6.2) in such a manner that the expression (6.2) does indeed satisfy the Equation (6.1).

But under what conditions is this assumption actually valid? That is, under what conditions can we be certain that the differential equation (6.1) actually *does* have a solution of the form (6.2)? This is a question of considerable importance; for it would be quite absurd to actually try to find a "solution" of the form (6.2) if there were really no such solution to be found! In order to answer this important question concerning the existence of a solution of the form (6.2), we shall first introduce certain basic definitions. For this purpose let us write the differential equation (6.1) in the equivalent normalized form

$$\frac{d^2y}{dx^2} + P_1(x)\frac{dy}{dx} + P_2(x)y = 0, \tag{6.3}$$

where

$$P_1(x) = \frac{a_1(x)}{a_0(x)} \quad \text{and} \quad P_2(x) = \frac{a_2(x)}{a_0(x)}.$$

DEFINITION

A function f is said to be analytic *at x_0 if its Taylor series about x_0,*

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

exists and converges to $f(x)$ for all x in some open interval including x_0.

We note that all polynomial functions are analytic everywhere; so also are the functions with values e^x, $\sin x$, and $\cos x$. A rational function is analytic except at those values of x at which its denominator is zero. For example, the rational function defined by $1/(x^2 - 3x + 2)$ is analytic except at $x = 1$ and $x = 2$.

DEFINITION

The point x_0 is called an ordinary point *of the differential equation (6.1) if* both *of the functions P_1 and P_2 in the equivalent normalized equation (6.3) are analytic at x_0. If either (or both) of these functions is* not *anaytic at x_0, then x_0 is called a* singular point *of the differential equation (6.1).*

THEOREM 6.1

Hypothesis. *The point x_0 is an ordinary point of the differential equation (6.1).*

Conclusion. *The differential equation (6.1) has two nontrivial linearly independent power series solutions of the form*

$$\sum_{n=0}^{\infty} c_n(x - x_0)^n, \tag{6.2}$$

and these power series converge in some interval $|x - x_0| < R$ (where $R > 0$) about x_0.

This theorem gives us a sufficient condition for the existence of power series solutions of the differential equation (6.1). It states that if x_0 is an ordinary point of equation (6.1), then this equation has *two* power series solutions in powers of $x - x_0$ and that these two power series solutions are *linearly independent*. Thus if x_0 is an ordinary point of (6.1), we may obtain the *general solution* of (6.1) as a linear combination of these two linearly independent power series. We shall omit the proof of this important theorem.

6.2 SOLUTIONS ABOUT SINGULAR POINTS; THE METHOD OF FROBENIUS

A. Regular Singular Points

We again consider the homogeneous linear differential equation

$$a_0(x)\frac{d^2y}{dx^2} + a_1(x)\frac{dy}{dx} + a_2(x)y = 0, \tag{6.1}$$

and we assume that x_0 is a *singular* point of (6.1). Then Theorem 6.1 does *not* apply at the point x_0, and we are *not* assured of a power series solution

$$y = \sum_{n=0}^{\infty} c_n(x - x_0)^n \tag{6.2}$$

of (6.1) in powers of $x - x_0$. Indeed an equation of the form (6.1) with a singular point at x_0 does *not*, in general, have a solution of the form (6.2). Clearly we must seek a different type of solution in such a case, but what type of solution can we expect? It happens that under certain conditions we are justified in assuming a solution of the form

$$y = |x - x_0|^r \sum_{n=0}^{\infty} c_n(x - x_0)^n, \tag{6.49}$$

where r is a certain (real or complex) constant. Such a solution is clearly a power series in $x - x_0$ multiplied by a certain *power* of $|x - x_0|$. In order to state conditions under which a solution of this form is assured, we proceed to classify singular points.

We again write the differential equation (6.1) in the equivalent normalized form

$$\frac{d^2y}{dx^2} + P_1(x)\frac{dy}{dx} + P_2(x)y = 0, \tag{6.3}$$

where

$$P_1(x) = \frac{a_1(x)}{a_0(x)} \quad \text{and} \quad P_2(x) = \frac{a_2(x)}{a_0(x)}.$$

DEFINITION

Consider the differential equation (6.1), and assume that at least one of the functions P_1 and P_2 in the equivalent normalized equation (6.3) is not *analytic at x_0, so that x_0 is a singular point of (6.1). If the functions defined by the products*

$$(x - x_0)P_1(x) \quad \textit{and} \quad (x - x_0)^2 P_2(x) \tag{6.50}$$

are both analytic at x_0, then x_0 is called a regular singular point *of the differential equation (6.1). If either (or both) of the functions defined by the products (6.50) is not analytic at x_0, then x_0 is called an* irregular singular point *of (6.1).*

THEOREM 6.2

Hypothesis. *The point x_0 is a regular singular point of the differential equation (6.1).*

Conclusion. *The differential equation (6.1) has* at least one *nontrivial solution of the form*

$$|x - x_0|^r \sum_{n=0}^{\infty} c_n(x - x_0)^n, \tag{6.49}$$

where r is a definite (real or complex) constant which may be determined, and this solution is valid in some deleted interval $0 < |x - x_0| < R$ (where $R > 0$) about x_0.

B. The Method of Frobenius

Now that we are assured of at least one solution of the form (6.49) about a regular singular point x_0 of the differential equation (6.1), how do we proceed to determine the coefficients c_n and the number r in this solution? The procedure is similar to that introduced in Section 6.1 and is commonly called the *method of Frobenius.* We shall briefly outline the method and then illustrate it by applying it to the differential equation (6.51). In this outline and the illustrative example that follows we shall seek solutions valid in some interval $0 < x - x_0 < R$. Note that for all such x, $|x - x_0|$ is simply $x - x_0$. To obtain solutions valid for $-R < x - x_0 < 0$, simply replace $x - x_0$ by $-(x - x_0) > 0$ and proceed as in the outline.

THEOREM 6.3

Hypothesis. *Let the point x_0 be a regular singular point of the differential equation (6.1). Let r_1 and r_2 [where $\mathrm{Re}(r_1) \geq \mathrm{Re}(r_2)$] be the roots of the indicial equation associated with x_0.*

Conclusion 1. *Suppose $r_1 - r_2 \neq N$, where N is a nonnegative integer (that is, $r_1 - r_2 \neq 0, 1, 2, 3, \ldots$). Then the differential equation (6.1) has two nontrivial linearly independent solutions y_1 and y_2 of the form (6.49) given respectively by*

$$y_1(x) = |x - x_0|^{r_1} \sum_{n=0}^{\infty} c_n (x - x_0)^n, \tag{6.65}$$

where $c_0 \neq 0$, and

$$y_2(x) = |x - x_0|^{r_2} \sum_{n=0}^{\infty} c_n^* (x - x_0)^n, \tag{6.66}$$

where $c_0^* \neq 0$.

Conclusion 2. *Suppose $r_1 - r_2 = N$, where N is a positive integer. Then the differential equation (6.1) has two nontrivial linearly independent solutions y_1 and y_2 given respectively by*

$$y_1(x) = |x - x_0|^{r_1} \sum_{n=0}^{\infty} c_n (x - x_0)^n, \tag{6.65}$$

where $c_0 \neq 0$, and

$$y_2(x) = |x - x_0|^{r_2} \sum_{n=0}^{\infty} c_n^* (x - x_0)^n + C y_1(x) \ln |x - x_0|, \tag{6.67}$$

where $c_0^* \neq 0$ and C is a constant which may or may not be zero.

Conclusion 3. *Suppose $r_1 - r_2 = 0$. Then the differential equation (6.1) has two nontrivial linearly independent solutions y_1 and y_2 given respectively by*

$$y_1(x) = |x - x_0|^{r_1} \sum_{n=0}^{\infty} c_n(x - x_0)^n, \tag{6.65}$$

where $c_0 \neq 0$, and

$$y_2(x) = |x - x_0|^{r_1+1} \sum_{n=0}^{\infty} c_n^*(x - x_0)^n + y_1(x)\ln|x - x_0|. \tag{6.68}$$

The solutions in Conclusions 1, 2, and 3 are valid in some deleted interval $0 < |x - x_0| < R$ about x_0.

In the illustrative examples and exercises that follow, we shall again seek solutions valid in some interval $0 < x - x_0 < R$. We shall therefore discuss the conclusions of Theorem 6.3 for such an interval. Before doing so, we again note that if $0 < x - x_0 < R$, then $|x - x_0|$ is simply $x - x_0$.

From the three conclusions of Theorem 6.3 we see that if x_0 is a regular singular point of (6.1), and $0 < x - x_0 < R$, then there is *always* a solution

$$y_1(x) = (x - x_0)^{r_1} \sum_{n=0}^{\infty} c_n(x - x_0)^n$$

of the form (6.49) for $0 < x - x_0 < R$ corresponding to the root r_1 of the indicial equation associated with x_0. Note again that the root r_1 is the *larger* root if r_1 and r_2 are real and unequal. From Conclusion 1 we see that if $0 < x - x_0 < R$ and the difference $r_1 - r_2$ between the roots of the indicial equation is *not* zero or a positive integer, then there is always a linearly independent solution

$$y_2(x) = (x - x_0)^{r_2} \sum_{n=0}^{\infty} c_n^*(x - x_0)^n$$

of the form (6.49) for $0 < x - x_0 < R$ corresponding to the root r_2. Note that the root r_2 is the *smaller* root if r_1 and r_2 are real and unequal. In particular, observe that if r_1 and r_2 are conjugate complex, then $r_1 - r_2$ is pure imaginary, and there will *always* be a linearly independent solution of the form (6.49) corresponding to r_2. However, from Conclusion 2 we see that if $0 < x - x_0 < R$ and the difference $r_1 - r_2$ *is a positive integer*, then a solution that is linearly independent of the "basic" solution of the form (6.49) for $0 < x - x_0 < R$ is of the generally more complicated form

$$y_2(x) = (x - x_0)^{r_2} \sum_{n=0}^{\infty} c_n^*(x - x_0)^n + Cy_1(x)\ln|x - x_0|$$

for $0 < x - x_0 < R$. Of course, if the constant C in this solution is zero, then it reduces to the simpler type of the form (6.49) for $0 < x - x_0 < R$. Finally, from Conclusion 3, we see that if $r_1 - r_2$ *is zero*, then the linearly independent solution $y_2(x)$ *always* involves the logarithmic term $y_1(x)\ln|x - x_0|$ and is *never* of the simple form (6.49) for $0 < x - x_0 < R$.

We shall now consider several examples that will (1) give further practice in the method of Frobenius, (2) illustrate the conclusions of Theorem 6.3, and (3) indicate how a linearly independent solution of the more complicated form involving the logarithmic term may be found in cases in which it exists. In each example we shall take $x_0 = 0$ and seek solutions valid in some interval $0 < x < R$. Thus note that in each example $|x - x_0| = |x| = x$.

6.3 BESSEL'S EQUATION AND BESSEL FUNCTIONS

A. Bessel's Equation of Order Zero

The differential equation

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - p^2)y = 0, \tag{6.101}$$

where p is a parameter, is called *Bessel's equation of order p*. Any solution of Bessel's equation of order p is called a *Bessel function of order p*. Bessel's equation and Bessel functions occur in connection with many problems of physics and engineering, and there is an extensive literature dealing with the theory and application of this equation and its solutions.

If $p = 0$, Equation (6.101) is equivalent to the equation

$$x\frac{d^2y}{dx^2} + \frac{dy}{dx} + xy = 0, \tag{6.102}$$

which is called *Bessel's equation of order zero*. We shall seek solutions of this equation that are valid in an interval $0 < x < R$. We observe at once that $x = 0$ is a regular singular point of (6.102); and hence, since we seek solutions for $0 < x < R$, we assume a solution

$$y = \sum_{n=0}^{\infty} c_n x^{n+r}, \tag{6.103}$$

where $c_0 \neq 0$. Upon differentiating (6.103) twice and substituting into (6.102), we obtain

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

Simplifying, we write this in the form

$$\sum_{n=0}^{\infty} (n+r)^2 c_n x^{n+r-1} + \sum_{n=2}^{\infty} c_{n-2} x^{n+r-1} = 0$$

or

$$r^2 c_0 x^{r-1} + (1+r)^2 c_1 x^r + \sum_{n=2}^{\infty} [(n+r)^2 c_n + c_{n-2}] x^{n+r-1} = 0. \tag{6.104}$$

Equating to zero the coefficient of the lowest power of x in (6.104), we obtain the indicial equation $r^2 = 0$, which has equal roots $r_1 = r_2 = 0$. Equating to zero the coefficients of the higher powers of x in (6.104) we obtain

$$(1+r)^2 c_1 = 0 \tag{6.105}$$

and the recurrence formula

$$(n+r)^2 c_n + c_{n-2} = 0, \qquad n \geq 2. \tag{6.106}$$

Letting $r = 0$ in (6.105), we find at once that $c_1 = 0$. Letting $r = 0$ in (6.106) we obtain the recurrence formula in the form

$$n^2 c_n + c_{n-2} = 0, \qquad n \geq 2,$$

or

$$c_n = -\frac{c_{n-2}}{n^2}, \qquad n \geq 2.$$

From this we obtain successively

$$c_2 = -\frac{c_0}{2^2}, \qquad c_3 = -\frac{c_1}{3^2} = 0 \text{ (since } c_1 = 0), \qquad c_4 = -\frac{c_2}{4^2} = \frac{c_0}{2^2 \cdot 4^2}, \dots .$$

We note that all odd coefficients are zero and that the general even coefficient may be written

$$c_{2n} = \frac{(-1)^n c_0}{2^2 \cdot 4^2 \cdot 6^2 \cdots (2n)^2} = \frac{(-1)^n c_0}{(n!)^2 2^{2n}}, \qquad n \geq 1.$$

Letting $r = 0$ in (6.103) and using these values of c_{2n}, we obtain the solution $y = y_1(x)$, where

$$y_1(x) = c_0 \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}$$

If we set the arbitrary constant $c_0 = 1$, we obtain an important particular solution of Equation (6.102). This particular solution defines a function denoted by J_0 and called the *Bessel function of the first kind of order zero*. That is, the function J_0 is the particular solution of Equation (6.102) defined by

$$J_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}. \tag{6.107}$$

Writing out the first few terms of this series solution, we see that

$$\begin{aligned} J_0(x) &= 1 - \frac{1}{(1!)^2}\left(\frac{x}{2}\right)^2 + \frac{1}{(2!)^2}\left(\frac{x}{2}\right)^4 - \frac{1}{(3!)^2}\left(\frac{x}{2}\right)^6 + \cdots \\ &= 1 - \frac{x^2}{4} + \frac{x^4}{64} - \frac{x^6}{2304} + \cdots. \end{aligned} \tag{6.108}$$

Since the roots of the indicial equation are equal, we know from Theorem 6.3 that a solution of Equation (6.102) which is linearly independent of J_0 must be of the form

$$y = x \sum_{n=0}^{\infty} c_n^* x^n + J_0(x) \ln x,$$

for $0 < x < R$. Also, we know that such a linearly independent solution can be found by the method of reduction of order (Section 4.1). Indeed from Theorem 4.7 we know that this linearly independent solution y_2 is given by

$$y_2(x) = J_0(x) \int \frac{e^{-\int dx/x}}{[J_0(x)]^2}\, dx$$

and hence by

$$y_2(x) = J_0(x) \int \frac{dx}{x[J_0(x)]^2}.$$

From (6.108) we find that

$$[J_0(x)]^2 = 1 - \frac{x^2}{2} + \frac{3x^4}{32} - \frac{5x^6}{576} + \cdots$$

and hence

$$\frac{1}{[J_0(x)]^2} = 1 + \frac{x^2}{2} + \frac{5x^4}{32} + \frac{23x^6}{576} + \cdots.$$

Thus

$$\begin{aligned} y_2(x) &= J_0(x)\int\left(\frac{1}{x}+\frac{x}{2}+\frac{5x^3}{32}+\frac{23x^5}{576}+\cdots\right)dx \\ &= J_0(x)\left(\ln x+\frac{x^2}{4}+\frac{5x^4}{128}+\frac{23x^6}{3456}+\cdots\right) \\ &= J_0(x)\ln x+\left(1-\frac{x^2}{4}+\frac{x^4}{64}-\frac{x^6}{2304}+\cdots\right)\left(\frac{x^2}{4}+\frac{5x^4}{128}+\frac{23x^6}{3456}+\cdots\right) \\ &= J_0(x)\ln x+\frac{x^2}{4}-\frac{3x^4}{128}+\frac{11x^6}{13824}+\cdots. \end{aligned}$$

We thus obtain the first few terms of the "second" solution y_2 by the method of reduction of order. However, our computations give no information concerning the general coefficient c_{2n}^* in the above series. Indeed, it seems unlikely that an expression for the general coefficient can be found. However, let us observe that

$$\begin{aligned} (-1)^2\frac{1}{2^2(1!)^2}(1) &= \frac{1}{2^2}=\frac{1}{4}, \\ (-1)^3\frac{1}{2^4(2!)^2}\left(1+\frac{1}{2}\right) &= -\frac{3}{2^4\cdot 2^2\cdot 2}=-\frac{3}{128}, \\ (-1)^4\frac{1}{2^6(3!)^2}\left(1+\frac{1}{2}+\frac{1}{3}\right) &= \frac{11}{2^6\cdot 6^2\cdot 6}=\frac{11}{13824}. \end{aligned}$$

Having observed these relations, we may express the solution y_2 in the following more systematic form:

$$y_2(x)=J_0(x)\ln x+\frac{x^2}{2^2}-\frac{x^4}{2^4(2!)^2}\left(1+\frac{1}{2}\right)+\frac{x^6}{2^6(3!)^2}\left(1+\frac{1}{2}+\frac{1}{3}\right)+\cdots.$$

Further, we would certainly suspect that the general coefficient c_{2n}^* is given by

$$c_{2n}^*=\frac{(-1)^{n+1}}{2^{2n}(n!)^2}\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}\right),\qquad n\geq 1.$$

It may be shown (though not without some difficulty) that this is indeed the case. This being true, we may express y_2 in the form

$$y_2(x)=J_0(x)\ln x+\sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^{2n}}{2^{2n}(n!)^2}\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}\right). \qquad (6.109)$$

Since the solution y_2 defined by (6.109) is linearly independent of J_0 we could write the general solution of the differential equation (6.102) as a general linear combination of J_0 and y_2. However, this is not usually done; instead, it has been customary to choose a certain special linear combination of J_0 and y_2 and take this special combination as the "second" solution of Equation (6.102). This special combination is defined by

$$\frac{2}{\pi}\left[y_2(x)+(\gamma-\ln 2)J_0(x)\right],$$

where γ is a number called *Euler's constant* and is defined by

$$\gamma = \lim_{n \to \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n\right) \approx 0.5772.$$

It is called the *Bessel function of the second kind of order zero* (Weber's form) and is commonly denoted by Y_0. Thus the second solution of (6.102) is commonly taken as the function Y_0, where

$$Y_0(x) = \frac{2}{\pi}\left[J_0(x)\ln x + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^{2n}}{2^{2n}(n!)^2}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) + (\gamma - \ln 2)J_0(x)\right]$$

or

$$Y_0(x) = \frac{2}{\pi}\left[\left(\ln \frac{x}{2} + \gamma\right)J_0(x) + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^{2n}}{2^{2n}(n!)^2}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right)\right]. \quad (6.110)$$

Therefore if we choose Y_0 as the second solution of the differential equation (6.102), the general solution of (6.102) for $0 < x < R$ is given by

$$y = c_1 J_0(x) + c_2 Y_0(x), \quad (6.111)$$

where c_1 and c_2 are arbitrary constants, and J_0 and Y_0 are defined by (6.107) and (6.110), respectively.

The functions J_0 and Y_0 have been studied extensively and tabulated. Many of the interesting properties of these functions are indicated by their graphs, which are shown in Figure 6.1.

9.1 DEFINITION, EXISTENCE, AND BASIC PROPERTIES OF THE LAPLACE TRANSFORM

A. Definition and Existence

DEFINITION

Let f be a real-valued function of the real variable t, defined for $t > 0$. Let s be a variable that we shall assume to be real, and consider the function F defined by

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt, \tag{9.1}$$

for all values of s for which this integral exists. The function F defined by the integral (9.1) is called the Laplace transform *of the function f. We shall denote the Laplace transform F of f by $\mathscr{L}\{f\}$ and shall denote $F(s)$ by $\mathscr{L}\{f(t)\}$.*

In order to be certain that the integral (9.1) does exist for some range of values of s, we must impose suitable restrictions upon the function f under consideration. We shall do this shortly; however, first we shall directly determine the Laplace transforms of a few simple functions.

9.3 LAPLACE TRANSFORM SOLUTION OF LINEAR DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

A. The Method

We now consider how the Laplace transform may be applied to solve the initial-value problem consisting of the nth-order linear differential equation with constant coefficients

$$a_0 \frac{d^n y}{dt^n} + a_1 \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_{n-1} \frac{dy}{dt} + a_n y = b(t), \tag{9.48}$$

plus the initial conditions

$$y(0) = c_0,\ y'(0) = c_1, \ldots, y^{(n-1)}(0) = c_{n-1}. \tag{9.49}$$

Theorem 4.1 (Chapter 4) assures us that this problem has a unique solution.

We now take the Laplace transform of both members of Equation (9.48). By Theorem 9.2, we have

$$a_0 \mathscr{L}\left\{\frac{d^n y}{dt^n}\right\} + a_1 \mathscr{L}\left\{\frac{d^{n-1} y}{dt^{n-1}}\right\} + \cdots + a_{n-1} \mathscr{L}\left\{\frac{dy}{dt}\right\} + a_n \mathscr{L}\{y(t)\} = \mathscr{L}\{b(t)\}. \tag{9.50}$$

We now apply Theorem 9.4 to

$$\mathscr{L}\left\{\frac{d^n y}{dt^n}\right\}, \mathscr{L}\left\{\frac{d^{n-1} y}{dt^{n-1}}\right\}, \ldots, \mathscr{L}\left\{\frac{dy}{dt}\right\}$$

in the left member of Equation (9.50). Using the initial conditions (9.49), we have

$$\begin{aligned}
\mathscr{L}\left\{\frac{d^n y}{dt^n}\right\} &= s^n \mathscr{L}\{y(t)\} - s^{n-1} y(0) - s^{n-2} y'(0) - \cdots - y^{(n-1)}(0) \\
&= s^n \mathscr{L}\{y(t)\} - c_0 s^{n-1} - c_1 s^{n-2} - \cdots - c_{n-1}, \\
\mathscr{L}\left\{\frac{d^{n-1}}{dt^{n-1}}\right\} &= s^{n-1} \mathscr{L}\{y(t)\} - s^{n-2} y(0) - s^{n-3} y'(0) - \cdots - y^{(n-2)}(0) \\
&= s^{n-1} \mathscr{L}\{y(t)\} - c_0 s^{n-2} - c_1 s^{n-3} - \cdots - c_{n-2}, \\
&\ \vdots \\
\mathscr{L}\left\{\frac{dy}{dt}\right\} &= s\mathscr{L}\{y(t)\} - y(0) = s\mathscr{L}\{y(t)\} - c_0.
\end{aligned}$$

Thus, letting $Y(s)$ denote $\mathscr{L}\{y(t)\}$ and $B(s)$ denote $\mathscr{L}\{b(t)\}$, Equation (9.50) becomes

$$\begin{aligned}
&[a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n] Y(s) \\
&\qquad - c_0[a_0 s^{n-1} + a_1 s^{n-2} + \cdots + a_{n-1}] \\
&\qquad - c_1[a_0 s^{n-2} + a_1 s^{n-3} + \cdots + a_{n-2}] \\
&\qquad - \cdots - c_{n-2}[a_0 s + a_1] - c_{n-1} a_0 = B(s).
\end{aligned} \tag{9.51}$$

Since b is a known function of t, then B, assuming it exists and can be determined, is a known function of s. Thus Equation (9.51) is an algebraic equation in the "unknown" $Y(s)$. We now solve the algebraic equation (9.51) to determine $Y(s)$. Once $Y(s)$ has been found, we then find the unique solution

$$y(t) = \mathscr{L}^{-1}\{Y(s)\}$$

of the given initial-value problem using the table of transforms.

We summarize this procedure as follows:

1. Take the Laplace transform of both sides of the differential equation (9.48), applying Theorem 9.4 and using the initial conditions (9.49) in the process, and equate the results to obtain the algebraic equation (9.51) in the "unknown" $Y(s)$.
2. Solve the algebraic equation (9.51) thus obtained to determine $Y(s)$.
3. Having found $Y(s)$, employ the table of transforms to determine the solution $y(t) = \mathscr{L}^{-1}\{Y(s)\}$ of the given initial-value problem.

Questions	opt1	opt2	opt3	opt4	opt 5	opt 6	Answer
Consider the equation $c_o(t)x''+c_1(t)x=0$ then a point a is an ordinary point if $d_1(t)$ and $d_2(t)$ are analytic at ___________	t=0	t=a	t=1	$t=a^2$			t=a
An hermite equation has an ordinary point at______	t=0	t=a	t=1	$t=a^2$			t=0
An analytic function for an hermite equation at t=0 is _______	–t and 1	t and 2	-2t and 2	2t and 1			-2t and 2
The legendre equation of order p is_________	(1-t) x″-2tx′+p(p+1)x=0	(1-t) x″-2tx′+(p+1)x=0	t2 x″-2tx′+p(p+1)x=0	(1-t2) x″-2tx′+p(p+1)x=0			(1-t2) x″-2tx′+p(p+1)x=0
When $p_n(t)$ is called an legendre polynomial?	Pn(1)=0	Pn(0)=1	Pn(1)=1	Pn(t)=1			Pn(1)=1
If $p_n(t)$ is a legendre polynomial then $_{-1}\int^1$ pn(t)dt=___________________	1/(2n+1)!	2/(2n+1)!	2/(n+1)!	1/(n+2)!			2/(2n+1)!
If $p_m(t)$ and $p_n(t)$ are legendre polynomials then $_{-1}\int^1$pn(t) pm(t)dt=______ if $m\neq n$	1	-1	2	0			0
If $p_n(t)$ is a legendre polynomial then $p_n(-1)$= 1 if n is______	Negative	odd	Even	positive			odd
The Bessel equation of order p is______	$t^2x''+tx+tx'+(t^2-p^2)x=0$	tx″+(1-t)x″+px=0	2x″+(1-t)x′+(1-p2)x=0	$tx''+(1-t)x+p^2x=0$			$t^2x''+tx+tx'+(t^2-p^2)x=0$
The Bessel function of the first kind d/dt (tpJp(t))=_________	$t^{-p}J_p(t)$	$t^pJ_{p-1}(t)$	$t^{-p}J_{p+1}(t)$	$t^pJ_p(t)$			$t^pJ_{p-1}(t)$
If $p_n(t)$ is the generating function then $p_n(-1)$=______	-1	0	$(1)^n$	$(-1)^n$			$(-1)^n$
The hermite equation is________	t2x″-2tx′+x=0	x″+tx′-2x=0	x″-2tx′+2x=0	tx″-tx′+x=0			x″-2tx′+2x=0

Questions	opt1	opt2	opt3	opt4	opt 5	opt 6	Answer
The legendre polynomial $p_n(t)$ can be express as________	$1/2^n n!\ D^n(t^2-1)^n$	$1/2^n n!\ D^n(t^2-1)^{n-1}$	$1/n!\ D^n(t^2-1)^n$	$1/2^n n!\ D^n(t^2-1)$			$1/2^n n!\ D^n(t^2-1)^n$
The order of equation is $(D^2+2D-8)y=0$ is______	1	2	0	8			2
The solution of ordinary differential equation of n order contains ___________ arbitrary constants	More than n	no	n	Atleast n			n
The n^{th} order ordinary linear homogeneous differential equation have_______	(n-1) singular solution	one singular solution	n-singular solution	no singular solution			no singular solution
The linearity principle for ordinary differential equation holds for_______	Non-homogeneous equation	linear differential equation	Homogeneous equation	non-linear equation			linear differential equation
A singular point which in______ is called an irregular singular point	Regular	ordinary point	analytic point	analytic function			Regular
If $p_m(t)$ and $p_n(t)$ are legendre polynomials then $_{-1}\int^1 pn(t)$ pm(t)dt=______ if m=n	0	1/n+1	2/(2n+1)	1			2/(2n+1)
On Bessel's function, where n is any integer then J-n(x)=________	$(-1)^n J_{-n}(x)$	$(-1)^n J_n(x)$	$(-1)^n J_{n+1}(x)$	$(-1)^n J_{n-1}(x)$			$(-1)^n J_n(x)$
When the hermite equation has an ordinary point?	t=0	t=-2	t=0	t=0			t=0
The second order linear homogeneous equation is of the form________	x″+a1(t)x′+a2(t)x	x″+a1(t)x′+a2(t)x= constant	x″+a1(x)x=0	x″+a1(x)x′=constant			x″+a1(t)x′+a2(t)x
The regular singular point of the equation tx″+(1-t)x′+nx=0 is_______	t=1	t=-1	t=0	t=n			t=0
The equation tx″+(1-t)x′+nx=0 where n is a constant, is called the_____________	aagrange equation	legendre equation	Bessel equation	hermite equation			lagrange equation

Questions	opt1	opt2	opt3	opt4	opt 5	opt 6	Answer
The singular point of the equation $t(t-1)^2(t+3)x''+t^2x'-(t^2+t-1)x=$ is___________	t=0 and =1	t=0, t=1 and t=-3	t=1 qnd t=-3	t=0 and t=-3			t=0, t=1 and t=-3
The equation $t^2x''-(1+t)x=0$ having a regular singular point at_______	t=-1	t=1	$t=\sqrt{-1}$	t=0			t=0
If Jp(t) is a Bessel function then d/dx[t-pJp(t)]=_________	$-t^pJ_{p-1}(t)$	$t^{-p}J_{p+1}(t)$	$-t^{-p}J_{p+1}(t)$	$t^pJ_{p-1}(t)$			$-t^{-p}J_{p+1}(t)$
The regular singular point of the equation $t^2\ x''+2tx'-n(n+1)x=$ is_________	0	infinity	1	2			infinity
The Bessel equation is of the second order then it possesses two _______	linearly dependent solution	independent solutions	dependent solutions	linearly independent solutions			linearly independent solutions
A point to is defined to be a singular point for the equations $a0(t)x''+a1(t)x'+a2(t)x=0$ if it is______	not an ordinary point	ordinary point	not an irregular point	irregular point			not an ordinary point
The regular singular points of the equations $(t-t^2)x''+[\gamma-(\alpha+\beta+1)]tx-\beta\alpha x=0$ is______________________	0and 1	0 and ∞	0,1 and ∞	1 and ∞			0,1 and ∞
The Bessel function of	(1/π)Jn(t)	$\pi J_n(t)$	π/Jn (t)	Jn(t)			$\pi J_n(t)$
The consider non-linear differential equation $x'=t^2-x^2$, x=1/2 when t=0 then the value of $x'(0)$=_____________	1/2	- 1/2	1/4	-1/4			-1/4
The equation $(1-t^2)x''-2tx'+p(p+1)x=0$ where p is a real number is called the ______of order p	legendre equation	laguerse equation	Bessel equation	Hermite equation			legendre equation
The Bessel equation possesses a ________ at t=0	ordinary point	analytic function	regular singular point	singular point			regular singular point
The equation $t(t-1)^2(t+3)x''+t^2x'-(t^2+t-1)=0$ is not analytic at ________	t=0	t=-1	t=-3	t=1			t=1

Questions	opt1	opt2	opt3	opt4	opt 5	opt 6	Answer
The Bessel function when n is _______	even or odd	odd	costant	even			odd
A regular singular point of the equation $2t^2x''+(2t+1)x'-x=0$ is ___________	t=0	t=2	t=1	t=-1			t=0
An _________ equation has an ordinary point at t = 0.	Legendre	Bessel	Hermite	Lagrange			Hermite
The _________ order linear homogeneous equation is of the form $x'' + a1(t)x'+a2(t)x= 0$	first	second	third	fourth			second

Simulation Modeling

Introduction

In many situations a modeler is unable to construct an analytic (symbolic) model adequately explaining the behavior being observed because of its complexity or the intractability of the proposed explicative model. Yet if it is necessary to make predictions about the behavior, the modeler may conduct experiments (or gather data) to investigate the relationship between the dependent variable(s) and selected values of the independent variable(s) within some range. We constructed empirical models based on collected data in Chapter 4. To collect the data the modeler may observe the behavior directly. In other instances, the behavior might be duplicated (possibly in a scaled-down version) under controlled conditions, as we will do when predicting the size of craters in Section 14.4.

In some circumstances, it may not be feasible either to observe the behavior directly or to conduct experiments. For instance, consider the service provided by a system of elevators during morning rush hour. After identifying an appropriate problem and defining what is meant by good service, we might suggest some alternative delivery schemes, such as assigning elevators to even and odd floors or using express elevators. Theoretically, each alternative could be tested for some period of time to determine which one provided the best service for particular arrival and destination patterns of the customers. However, such a procedure would probably be very disruptive because it would be necessary to harass the customers constantly as the required statistics were collected. Moreover, the customers would become very confused because the elevator delivery system would keep changing. Another problem concerns testing alternative schemes for controlling automobile traffic in a large city. It would be impractical to constantly change directions of the one-way streets and the distribution of traffic signals to conduct tests.

In still other situations, the system for which alternative procedures need to be tested *may not even exist yet*. An example is the situation of several proposed communications networks, with the problem of determining which is best for a given office building. Still another example is the problem of determining locations of machines in a new industrial plant. The *cost* of conducting experiments may be prohibitive. This is the case when an agency tries to predict the effects of various alternatives for protecting and evacuating the population in case of failure of a nuclear power plant.

In cases where the behavior cannot be explained analytically or data collected directly, the modeler might *simulate* the behavior indirectly in some manner and then test the various alternatives under consideration to estimate how each affects the behavior. Data can then be collected to determine which alternative is best. An example is to determine the drag force on a proposed submarine. Because it is infeasible to build a prototype, we can build a scaled model to simulate the behavior of the actual submarine. Another example of this type of simulation is using a scaled model of a jet airplane in a wind tunnel to estimate the effects of very high speeds for various designs of the aircraft. There is yet another type of simulation, which we will study in this chapter. This **Monte Carlo simulation** is typically accomplished with the aid of a computer.

Suppose we are investigating the service provided by a system of elevators at morning rush hour. In Monte Carlo simulation, the arrival of customers at the elevators during the hour and the destination floors they select need to be replicated. That is, the distribution of arrival times and the distribution of floors desired on the simulated trial must portray a possible rush hour. Moreover, after we have simulated many trials, the daily distribution of arrivals and destinations that occur must mimic the real-world distributions in proper proportions. When we are satisfied that the behavior is adequately duplicated, we can investigate various alternative strategies for operating the elevators. Using a large number of trials, we can gather appropriate statistics, such as the average total delivery time of a customer or the length of the longest queue. These statistics can help determine the best strategy for operating the elevator system.

This chapter provides a brief introduction to Monte Carlo simulation. Additional studies in probability and statistics are required to delve into the intricacies of computer simulation and understand its appropriate uses. Nevertheless, you will gain some appreciation of this powerful component of mathematical modeling. Keep in mind that there is a danger in placing too much confidence in the predictions resulting from a simulation, especially if the assumptions inherent in the simulation are not clearly stated. Moreover, the appearance of using large amounts of data and huge amounts of computer time, coupled with the fact the lay people can understand a simulation model and computer output with relative ease, often leads to overconfidence in the results.

When any Monte Carlo simulation is performed, random numbers are used. We discuss how to generate random numbers in Section 5.2. Loosely speaking, a "sequence of random numbers uniformly distributed in an interval m to n" is a set of numbers with no apparent pattern, where each number between m and n can appear with equal likelihood. For example, if you toss a six-sided die 100 times and write down the number showing on the die each time, you will have written down a sequence of 100 random integers approximately uniformly distributed over the interval 1 to 6. Now, suppose that random numbers consisting of six digits can be generated. The tossing of a coin can be duplicated by generating a random number and assigning it a head if the random number is even and a tail if the random number is odd. If this trial is replicated a large number of times, you would expect heads to occur about 50% of the time. However, there is an element of chance involved. It is possible that a run of 100 trials could produce 51 heads and that the next 10 trials could produce all heads (although this is not very likely). Thus, the estimate with 110 trials would actually be worse than the estimate with 100 trials. Processes with an element of chance involved are called **probabilistic**, as opposed to **deterministic**, processes. Monte Carlo simulation is therefore a probabilistic model.

The modeled behavior may be either deterministic or probabilistic. For instance, the area under a curve is deterministic (even though it may be impossible to find it precisely). On the other hand, the time between arrivals of customers at the elevator on a particular day is probabilistic behavior. Referring to Figure 5.1, we see that a deterministic model can be used to approximate either a deterministic or a probabilistic behavior, and likewise, a Monte Carlo simulation can be used to approximate a deterministic behavior (as you will see with

Figure 5.1
The behavior and the model can be either deterministic or probabilistic.

a Monte Carlo approximation to an area under a curve) or a probabilistic one. However, as we would expect, the real power of Monte Carlo simulation lies in modeling a probabilistic behavior.

A principal advantage of Monte Carlo simulation is the relative ease with which it can sometimes be used to approximate very complex probabilistic systems. Additionally, Monte Carlo simulation provides performance estimation over a wide range of conditions rather than a very restricted range as often required by an analytic model. Furthermore, because a particular submodel can be changed rather easily in a Monte Carlo simulation (such as the arrival and destination patterns of customers at the elevators), there is the potential of conducting a sensitivity analysis. Still another advantage is that the modeler has control over the level of detail in a simulation. For example, a very long time frame can be compressed or a small time frame expanded, giving a great advantage over experimental models. Finally, there are very powerful, high-level simulation languages (such as GPSS, GASP, PROLOG, SIMAN, SLAM, and DYNAMO) that eliminate much of the tedious labor in constructing a simulation model.

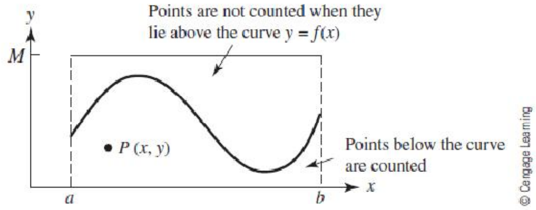
On the negative side, simulation models are typically expensive to develop and operate. They may require many hours to construct and large amounts of computer time and memory to run. Another disadvantage is that the probabilistic nature of the simulation model limits the conclusions that can be drawn from a particular run unless a sensitivity analysis is conducted. Such an analysis often requires many more runs just to consider a small number of combinations of conditions that can occur in the various submodels. This limitation then forces the modeler to estimate which combination might occur for a particular set of conditions.

Simulating Deterministic Behavior: Area Under a Curve

In this section we illustrate the use of Monte Carlo simulation to model a deterministic behavior, the area under a curve. We begin by finding an approximate value to the area under a nonnegative curve. Specifically, suppose $y = f(x)$ is some given continuous function satisfying $0 \le f(x) \le M$ over the closed interval $a \le x \le b$. Here, the number M is simply some constant that *bounds* the function. This situation is depicted in Figure 5.2. Notice that the area we seek is wholly contained within the rectangular region of height M and length $b - a$ (the length of the interval over which f is defined).

Now we select a point $P(x, y)$ at random from within the rectangular region. We will do so by generating two random numbers, x and y, satisfying $a \le x \le b$ and $0 \le y \le M$, and interpreting them as a point P with coordinates x and y. Once $P(x, y)$ is selected, we ask whether it lies within the region below the curve. That is, does the y-coordinate satisfy $0 \le y \le f(x)$? If the answer is yes, then count the point P by adding 1 to some counter.

■ **Figure 5.2**
The area under the nonnegative curve $y = f(x)$ over $a \le x \le b$ is contained within the rectangle of height M and base length $b - a$.



Two counters will be necessary: one to count the total points generated and a second to count those points that lie below the curve (Figure 5.2). You can then calculate an approximate value for the area under the curve by the following formula:

$$\frac{\text{area under curve}}{\text{area of rectangle}} \approx \frac{\text{number of points counted below curve}}{\text{total number of random points}}$$

As discussed in the Introduction, the Monte Carlo technique is probabilistic and typically requires a large number of trials before the deviation between the predicted and true values becomes small. A discussion of the number of trials needed to ensure a predetermined level of confidence in the final estimate requires a background in statistics. However, as a general rule, to double the accuracy of the result (i.e., to cut the expected error in half), about four times as many experiments are necessary.

The following algorithm gives the sequence of calculations needed for a general computer simulation of this Monte Carlo technique for finding the area under a curve.

Monte Carlo Area Algorithm

Input Total number n of random points to be generated in the simulation.

Output AREA = approximate area under the specified curve $y = f(x)$ over the given interval $a \le x \le b$, where $0 \le f(x) < M$.

Step 1 Initialize: COUNTER = 0.

Step 2 For $i = 1, 2, \ldots, n$, do Steps 3–5.

Step 3 Calculate random coordinates x_i and y_i that satisfy $a \le x_i \le b$ and $0 \le y_i < M$.

Step 4 Calculate $f(x_i)$ for the random x_i coordinate.

Step 5 If $y_i \le f(x_i)$, then increment the COUNTER by 1. Otherwise, leave COUNTER as is.

Step 6 Calculate AREA = $M(b - a)$ COUNTER/n.

Step 7 OUTPUT (AREA)

STOP

Table 5.1 gives the results of several different simulations to obtain the area beneath the curve $y = \cos x$ over the interval $-\pi/2 \le x \le \pi/2$, where $0 \le \cos x < 2$.

The actual area under the curve $y = \cos x$ over the given interval is 2 square units. Note that even with the relatively large number of points generated, the error is significant. For functions of one variable, the Monte Carlo technique is generally not competitive with quadrature techniques that you will learn in numerical analysis. The lack of an error bound and the difficulty in finding an upper bound M are disadvantages as well. Nevertheless, the

Table 5.1 Monte Carlo approximation to the area under the curve $y = \cos x$ over the interval $-\pi/2 \le x \le \pi/2$

Number of points	Approximation to area	Number of points	Approximation to area
100	2.07345	2000	1.94465
200	2.13628	3000	1.97711
300	2.01064	4000	1.99962
400	2.12058	5000	2.01429
500	2.04832	6000	2.02319
600	2.09440	8000	2.00669
700	2.02857	10000	2.00873
800	1.99491	15000	2.00978
900	1.99666	20000	2.01093
1000	1.96664	30000	2.01186

Monte Carlo technique can be extended to functions of several variables and becomes more practical in that situation.

Volume Under a Surface

Let's consider finding part of the volume of the sphere

$$x^2 + y^2 + z^2 \le 1$$

that lies in the first octant, $x > 0$, $y > 0$, $z > 0$ (Figure 5.3).

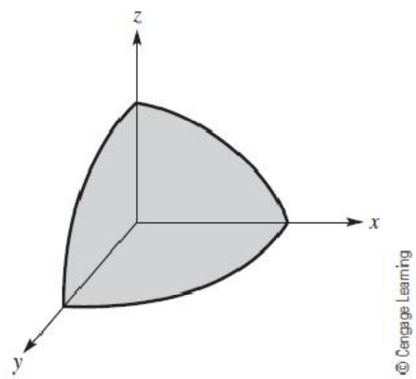
The methodology to approximate the volume is very similar to that of finding the area under a curve. However, now we will use an approximation for the volume under the surface by the following rule:

$$\frac{\text{volume under surface}}{\text{volume of box}} \approx \frac{\text{number of points counted below surface in 1st octant}}{\text{total number of points}}$$

The following algorithm gives the sequence of calculations required to employ Monte Carlo techniques to find the approximate volume of the region.

Figure 5.3

Volume of a sphere $x^2 + y^2 + z^2 \leq 1$ that lies in the first octant, $x > 0$, $y > 0$, $z > 0$



Monte Carlo Volume Algorithm

Input Total number n of random points to be generated in the simulation.

Output VOLUME = approximate volume enclosed by the specified function, $z = f(x, y)$ in the first octant, $x > 0$, $y > 0$, $z > 0$.

Step 1 Initialize: COUNTER = 0.

Step 2 For $i = 1, 2, \ldots, n$, do Steps 3–5.

Step 3 Calculate random coordinates x_i, y_i, z_i that satisfy $0 \le x_i \le 1$, $0 \le y_i \le 1$, $0 \le z_i \le 1$. (In general, $a \le x_i \le b$, $c \le y_i \le d$, $0 \le z_i \le M$.)

Step 4 Calculate $f(x_i, y_i)$ for the random coordinate (x_i, y_i).

Step 5 If random $z_i \le f(x_i, y_i)$, then increment the COUNTER by 1. Otherwise, leave COUNTER as is.

Step 6 Calculate VOLUME $= M(d - c)(b - a)$COUNTER$/n$.

Step 7 OUTPUT (VOLUME)

STOP

Table 5.2 gives the results of several Monte Carlo runs to obtain the approximate volume of

$$x^2 + y^2 + z^2 \le 1$$

that lies in the first octant, $x > 0$, $y > 0$, $z > 0$.

Table 5.2 Monte Carlo approximation to the volume in the first octant under the surface $x^2 + y^2 + z^2 \le 1$

Number of points	Approximate volume
100	0.4700
200	0.5950
300	0.5030
500	0.5140
1,000	0.5180
2,000	0.5120
5,000	0.5180
10,000	0.5234
20,000	0.5242

The actual volume in the first octant is found to be approximately 0.5236 cubic units $(\pi/6)$. Generally, though not uniformly, the error becomes smaller as the number of points generated increases.

Generating Random Numbers

In the previous section, we developed algorithms for Monte Carlo simulations to find areas and volumes. A key ingredient common to these algorithms is the need for random numbers. Random numbers have a variety of applications, including gambling problems, finding an area or volume, and modeling larger complex systems such as large-scale combat operations or air traffic control situations.

In some sense a computer does not really generate random numbers, because computers employ deterministic algorithms. However, we can generate sequences of pseudorandom numbers that, for all practical purposes, may be considered random. There is no single best random number generator or best test to ensure randomness.

There are complete courses of study for random numbers and simulations that cover in depth the methods and tests for pseudorandom number generators. Our purpose here is to introduce a few random number methods that can be utilized to generate sequences of numbers that are nearly random.

Many programming languages, such as Pascal and Basic, and other software (e.g., Minitab, MATLAB, and EXCEL) have built-in random number generators for user convenience.

Middle-Square Method

The middle-square method was developed in 1946 by John Von Neuman, S. Ulm, and N. Metropolis at Los Alamos Laboratories to simulate neutron collisions as part of the Manhattan Project. Their middle-square method works as follows:

1. Start with a four-digit number x_0, called the *seed*.
2. Square it to obtain an eight-digit number (add a leading zero if necessary).
3. Take the middle four digits as the next random number.

Continuing in this manner, we obtain a sequence that appears to be random over the integers from 0 to 9999. These integers can then be scaled to any interval a to b. For example, if we wanted numbers from 0 to 1, we would divide the four-digit numbers by 10,000. Let's illustrate the middle-square method.

Pick a seed, say $x_0 = 2041$, and square it (adding a leading zero) to get 04165681. The middle four digits give the next random number, 1656. Generating 13 random numbers in this way yields

n	0	1	2	3	4	5	6	7	8	9	10	11	12
x_n	2041	1656	7423	1009	0180	0324	1049	1004	80	64	40	16	2

We can use more than 4 digits if we wish, but we always take the middle number of digits equal to the number of digits in the seed. For example, if $x_0 = 653217$ (6 digits), its square 426,692,449,089 has 12 digits. Thus, take the middle 6 digits as the random number, namely, 692449.

The middle-square method is reasonable, but it has a major drawback in its tendency to degenerate to zero (where it will stay forever). With the seed 2041, the random sequence does seem to be approaching zero. How many numbers can be generated until we are almost at zero?

Linear Congruence

The linear congruence method was introduced by D. H. Lehmer in 1951, and a majority of pseudorandom numbers used today are based on this method. One advantage it has over other methods is that seeds can be selected that generate patterns that eventually cycle (we illustrate this concept with an example). However, the length of the cycle is so large that the pattern does not repeat itself on large computers for most applications. The method requires the choice of three integers: a, b, and c. Given some initial seed, say x_0, we generate a sequence by the rule

$$x_{n+1} = (a \times x_n + b) \bmod(c)$$

where c is the modulus, a is the multiplier, and b is the increment. The qualifier $\bmod(c)$ in the equation means to obtain the remainder after dividing the quantity $(a \times x_n + b)$ by c. For example, with $a = 1$, $b = 7$, and $c = 10$,

$$x_{n+1} = (1 \times x_n + 7) \bmod(10)$$

means x_{n+1} is the integer remainder upon dividing $x_n + 7$ by 10. Thus, if $x_n = 115$, then $x_{n+1} = \textit{remainder}\ (\frac{122}{10}) = 2$.

Before investigating the linear congruence methodology, we need to discuss **cycling**, which is a major problem that occurs with random numbers. Cycling means the sequence repeats itself, and, although undesirable, it is unavoidable. At some point, all pseudorandom number generators begin to cycle. Let's illustrate cycling with an example.

If we set our seed at $x_0 = 7$, we find $x_1 = (1 \times 7 + 7) \bmod(10)$ or $14 \bmod(10)$, which is 4. Repeating this same procedure, we obtain the sequence

$$7, 4, 1, 8, 5, 2, 9, 6, 3, 0, 7, 4, \ldots$$

and the original sequence repeats again and again. Note that there is cycling after 10 numbers. The methodology produces a sequence of integers between 0 and $c - 1$ inclusively before cycling (which includes the possible remainders after dividing the integers by c). Cycling is guaranteed with at most c numbers in the random number sequence. Nevertheless, c can be chosen to be very large, and a and b can be chosen in such a way as to obtain a full set of c numbers before cycling begins to occur. Many computers use $c = 2^{31}$ for the large value of c. Again, we can scale the random numbers to obtain a sequence between any limits a and b, as required.

A second problem that can occur with the linear congruence method is lack of statistical independence among the members in the list of random numbers. Any correlations between the nearest neighbors, the next-nearest neighbors, the third-nearest neighbors, and so forth are generally unacceptable. (Because we live in a three-dimensional world, third-nearest neighbor correlations can be particularly damaging in physical applications.) Pseudorandom number sequences can never be completely statistically independent because they are generated by a mathematical formula or algorithm. Nevertheless, the sequence will appear (for practical purposes) independent when it is subjected to certain statistical tests. These concerns are best addressed in a course in statistics.

Questions	opt1	opt2	opt3	opt4	opt5	opt6	Answer
When we finding an approximate value to the area under a curve will be ____	non negative	non positive	positive	neither negative nor positive			non positive
In Monte Carlo method functions are created by using a series of _____	positive numbers	negative numbers	random numbers	digital numbers			random numbers
The simulation model can be classified into ____ categories	4	3	5	2			4
Which models do not take variable time into consideration _____	static model	Monte Carlo model	probability model	deterministic model			static model
The area under a curve functions must be _____	continuous	modular	finite	infinite			continuous
Processes with an elt of chance involved are called _____	finite chance	deterministic	probabilistic	infinite chance			finite chance
Probabilistic as opposed to_____	deterministic	event	outcomes	event or outcome			deterministic
Monte Carlo simulation is therefore a _____model	random	deterministic	probabilistic	uniform			probabilistic
The area under a curve is _____model	deterministic	probabilistic	random	uniform			deterministic
Monte Carlo simulation can be used a _____ behaviour	select	approximate	predefined	choose			approximate
The real power of Monte Carlo simulation lies in modelling a ____behaviour	probabilistic	deterministic	random	probabilistic or deterministic			probabilistic
Example for higher level simulation languages is _____	SLAM , DYNAMO	Java	Pascal	Fortron77			SLAM , DYNAMO
Simulation models are typically _____ to develop and operate	low cost	very low cost	expansive	high cost			expansive
If y=f(x) is given continuous function satisfying $0 \leq f(x) \leq m$ over the closed interval _____	$a \leq x \leq b$	$a \leq x < b$	$a < x \leq b$	$a < x < b$			$a \leq x \leq b$

_____ technique is used to finding the area under curve	Monte Carlo	simulation	random selection	neither simulation nor random selection			Monte Carlo
If y=f(x) is given continuous function satisfying _____over the closed interval $a \le x \le b$	$0 \le f(x) \le m$	$0 \le f(x)<m$	$0<f(x) \le m$	$0<f(x)< m$			$0 \le f(x) \le m$
. _____ technique is used to find the volume under a surface	Monte Carlo	simulation	random selection	neither simulation nor random selection			Monte Carlo
Volume under surface is _____	m(d-c)(b-a) COUNTER/n	m(d-c)	m(a-b)/n	m(d-c)(b-a)/n			m(d-c)(b-a) COUNTER/n
Area under curve is ____	m(b-a) COUNTER/n	m(d-c)COUNTER/n	m(d-c)	m(b-a)/n			m(b-a) COUNTER/n
Middle square method starts with ____ digit number	four	two	eight	three			four
The middle square method was developed in _____	1947	1956	1946	1966			1946
Finding an air traffic control situation is one of the application of ___ numbers	selection	pseudo	random	middle square			pseudo
Pseudo random numbers based on ____ methods	middle square	random number	linear congruence	deterministic			random number
Middle square method starts with four digit number is called ____	leading seed	added seed	random seed	seed			seed
Depth of random number and simulation is ____	techniques	random number	pseudo random numbers	pseudo numbers			pseudo random numbers
Linear congruence depends on	random number	pseudo random numbers	cycling	neither cycling nor random numbers			cycling
_________ models are typically expansive to develop and operate	Simulation	very low cost	low cost	high cost			Simulation
_________ method starts with four digit number	Middle square	Linear congruence	Probabilistic	deterministic			Middle square

Queuing Models

A Harbor System

Consider a small harbor with unloading facilities for ships. Only one ship can be unloaded at any one time. Ships arrive for unloading of cargo at the harbor, and the time between the arrival of successive ships varies from 15 to 145 min. The unloading time required for a ship depends on the type and amount of cargo and varies from 45 to 90 min. We seek answers to the following questions:

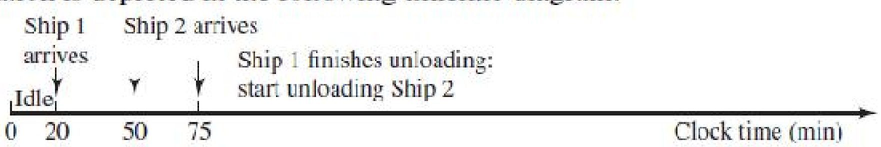
1. What are the average and maximum times per ship in the harbor?
2. If the *waiting time* for a ship is the time between its arrival and the start of unloading, what are the average and maximum waiting times per ship?
3. What percentage of the time are the unloading facilities idle?
4. What is the length of the longest queue?

To obtain some reasonable answers, we can simulate the activity in the harbor using a computer or programmable calculator. We assume the arrival times between successive ships and the unloading time per ship are uniformly distributed over their respective time intervals. For instance, the arrival time between ships can be any integer between 15 and 145, and any integer within that interval can appear with equal likelihood. Before giving a general algorithm to simulate the harbor system, let's consider a hypothetical situation with five ships.

We have the following data for each ship:

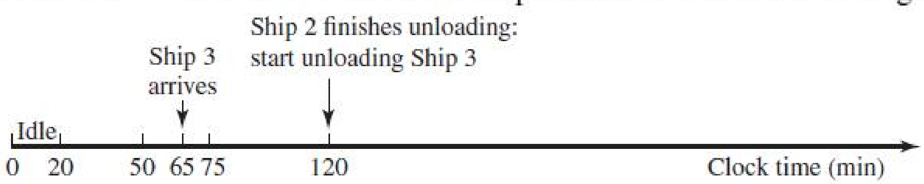
	Ship 1	Ship 2	Ship 3	Ship 4	Ship 5
Time between successive ships	20	30	15	120	25
Unloading time	55	45	60	75	80

Because Ship 1 arrives 20 min after the clock commences at $t = 0$ min, the harbor facilities are idle for 20 min at the start. Ship 1 immediately begins to unload. The unloading takes 55 min; meanwhile, Ship 2 arrives on the scene at $t = 20 + 30 = 50$ min after the clock begins. Ship 2 cannot start to unload until Ship 1 finishes unloading at $t = 20 + 55 = 75$ min. This means that Ship 2 must wait $75 - 50 = 25$ min before unloading begins. The situation is depicted in the following timeline diagram:



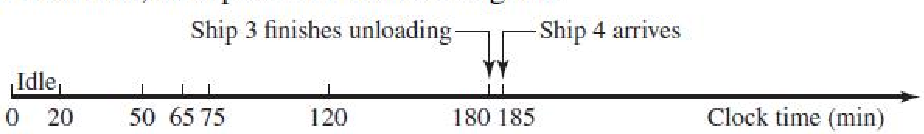
Timeline 1

Now before Ship 2 starts to unload, Ship 3 arrives at time $t = 50 + 15 = 65$ min. Becau the unloading of Ship 2 starts at $t = 75$ min and it takes 45 min to unload, unloading Sl 3 cannot start until $t = 75 + 45 = 120$ min, when Ship 2 is finished. Thus, Ship 3 m wait $120 - 65 = 55$ min. The situation is depicted in the next timeline diagram:



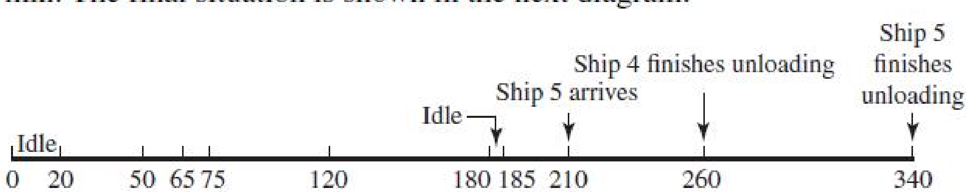
Timeline 2

Ship 4 does not arrive in the harbor until $t = 65 + 120 = 185$ min. Therefore, Sl 3 has already finished unloading at $t = 120 + 60 = 180$ min, and the harbor facilities idle for $185 - 180 = 5$ min. Moreover, the unloading of Ship 4 commences immediat upon its arrival, as depicted in the next diagram:



Timeline 3

Finally, Ship 5 arrives at $t = 185 + 25 = 210$ min, before Ship 4 finishes unloading $t = 185 + 75 = 260$ min. Thus, Ship 5 must wait $260 - 210 = 50$ min before it sta to unload. The simulation is complete when Ship 5 finishes unloading at $t = 260 + 80$ 340 min. The final situation is shown in the next diagram:



Timeline 4

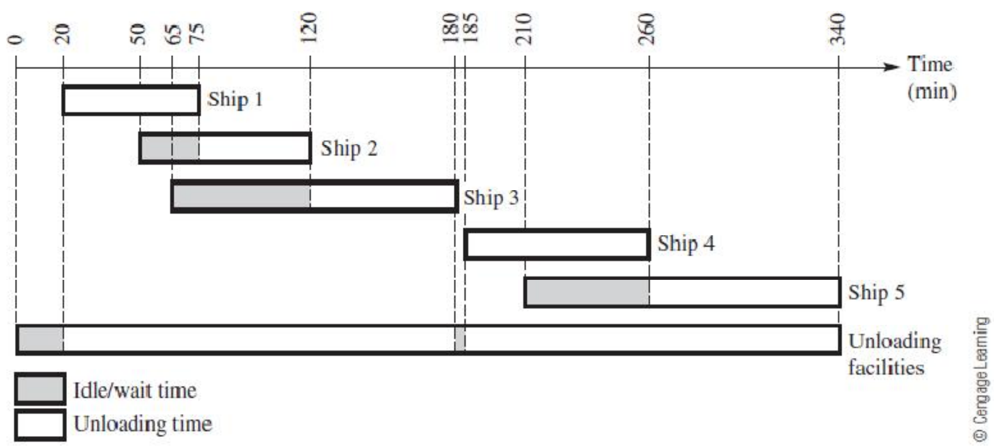


Figure 5.9
Idle and unloading times for the ships and docking facilities

In Figure 5.9, we summarize the waiting and unloading times for each of the five hypothetical ship arrivals. In Table 5.14, we summarize the results of the entire simulation of the five hypothetical ships. Note that the total waiting time spent by all five ships before unloading is 130 min. This waiting time represents a cost to the shipowners and is a source of customer dissatisfaction with the docking facilities. On the other hand, the docking facility has only 25 min of total idle time. It is in use 315 out of the total 340 min in the simulation, or approximately 93% of the time.

Suppose the owners of the docking facilities are concerned with the quality of service they are providing and want various management alternatives to be evaluated to determine whether improvement in service justifies the added cost. Several statistics can help in evaluating the quality of the service. For example, the maximum time a ship spends in the harbor is 130 min by Ship 5, whereas the average is 89 min (Table 5.14). Generally, customers are very sensitive to the amount of time spent waiting. In this example, the maximum time spent waiting for a facility is 55 min, whereas the average time spent waiting is 26 min. Some customers are apt to take their business elsewhere if queues are too long. In this case, the longest queue is two. The following Monte Carlo simulation algorithm computes such statistics to assess various management alternatives.

Table 5.14 Summary of the harbor system simulation

Ship no.	Random time between ship arrivals	Arrival time	Start service	Queue length at arrival	Wait time	Random unload time	Time in harbor	Dock idle time
1	20	20	20	0	0	55	55	20
2	30	50	75	1	25	45	70	0
3	15	65	120	2	55	60	115	0
4	120	185	185	0	0	75	75	5
5	25	210	260	1	50	80	130	0
Total (if appropriate):					130			25
Average (if appropriate):					26	63	89	

Summary of Harbor System Algorithm Terms

between$_i$ Time between successive arrivals of Ships i and $i - 1$ (a random integer varying between 15 and 145 min)
arrive$_i$ Time from start of clock at $t = 0$ when Ship i arrives at the harbor for unloading
unload$_i$ Time required to unload Ship i at the dock (a random integer varying between 45 and 90 min)
start$_i$ Time from start of clock at which Ship i commences its unloading
idle$_i$ Time for which dock facilities are idle immediately *before* commencement of unloading Ship i
wait$_i$ Time Ship i waits in the harbor after arrival before unloading commences
finish$_i$ Time from start of clock at which service for Ship i is completed at the unloading facilities
harbor$_i$ Total time Ship i spends in the harbor
HARTIME Average time per ship in the harbor
MAXHAR Maximum time of a ship in the harbor
WAITIME Average waiting time per ship before unloading
MAXWAIT Maximum waiting time of a ship
IDLETIME Percentage of total simulation time unloading facilities are idle

Harbor System Simulation Algorithm

Input Total number n of ships for the simulation.

Output HARTIME, MAXHAR, WAITIME, MAXWAIT, and IDLETIME.

Step 1 Randomly generate between_1 and unload_1. Then set $\text{arrive}_1 = \text{between}_1$.

Step 2 Initialize all output values:

$$\text{HARTIME} = \text{unload}_1, \qquad \text{MAXHAR} = \text{unload}_1,$$
$$\text{WAITIME} = 0, \qquad \text{MAXWAIT} = 0, \qquad \text{IDLETIME} = \text{arrive}_1$$

Step 3 Calculate finish time for unloading of Ship_1:

$$\text{finish}_1 = \text{arrive}_1 + \text{unload}_1$$

Step 4 For $i = 2, 3, \ldots, n$, do Steps 5–16.

Step 5 Generate the random pair of integers between_i and unload_i over their respective time intervals.

Step 6 Assuming the time clock begins at $t = 0$ min, calculate the time of arrival for Ship_i:

$$\text{arrive}_i = \text{arrive}_{i-1} + \text{between}_i$$

Step 7 Calculate the time difference between the arrival of Ship_i and the finish time for unloading the previous Ship_{i-1}:

$$\text{timediff} = \text{arrive}_i - \text{finish}_{i-1}$$

Step 8 For nonnegative timediff, the unloading facilities are idle:

$$\text{idle}_i = \text{timediff} \quad \text{and} \quad \text{wait}_i = 0$$

For negative timediff, Ship_i must wait before it can unload:

$$\text{wait}_i = -\text{timediff} \quad \text{and} \quad \text{idle}_i = 0$$

Step 9 Calculate the start time for unloading Ship_i:

$$\text{start}_i = \text{arrive}_i + \text{wait}_i$$

Step 10 Calculate the finish time for unloading Ship_i:

$$\text{finish}_i = \text{start}_i + \text{unload}_i$$

Step 11 Calculate the time in harbor for Ship_i:

$$\text{harbor}_i = \text{wait}_i + \text{unload}_i$$

Step 12 Sum harbor_i into total harbor time HARTIME for averaging.

Step 13 If $\text{harbor}_i > \text{MAXHAR}$, then set $\text{MAXHAR} = \text{harbor}_i$. Otherwise leave MAXHAR as is.

Step 14 Sum wait_i into total waiting time WAITIME for averaging.

Step 15 Sum idle_i into total idle time IDLETIME.

Step 16 If $\text{wait}_i > \text{MAXWAIT}$, then set $\text{MAXWAIT} = \text{wait}_i$. Otherwise leave MAXWAIT as is.

Step 17 Set $\text{HARTIME} = \text{HARTIME}/n$, $\text{WAITIME} = \text{WAITIME}/n$, and $\text{IDLETIME} = \text{IDLETIME}/\text{finish}_n$.

Step 18 OUTPUT (HARTIME, MAXHAR, WAITIME, MAXWAIT, IDLETIME)
STOP

Table 5.15 gives the results, according to the preceding algorithm, of six independent simulation runs of 100 ships each.

Now suppose you are a consultant for the owners of the docking facilities. What would be the effect of hiring additional labor or acquiring better equipment for unloading cargo so that the unloading time interval is reduced to between 35 and 75 min per ship? Table 5.16 gives the results based on our simulation algorithm.

You can see from Table 5.16 that a reduction of the unloading time per ship by 10 to 15 min decreases the time ships spend in the harbor, especially the waiting times. However, the percentage of the total time during which the dock facilities are idle nearly doubles. The situation is favorable for shipowners because it increases the availability of each ship for hauling cargo over the long run. Thus, the traffic coming into the harbor is likely to increase. If the traffic increases to the extent that the time between successive ships is reduced to between 10 and 120 min, the simulated results are as shown in Table 5.17. We can see from this table that the ships again spend more time in the harbor with the increased traffic, but now harbor facilities are idle much less of the time. Moreover, both the shipowners and the dock owners are benefiting from the increased business.

Table 5.15 Harbor system simulation results for 100 ships

Average time of a ship in the harbor	106	85	101	116	112	94
Maximum time of a ship in the harbor	287	180	233	280	234	264
Average waiting time of a ship	39	20	35	50	44	27
Maximum waiting time of a ship	213	118	172	203	167	184
Percentage of time dock facilities are idle	0.18	0.17	0.15	0.20	0.14	0.21

Note: All times are given in minutes. Time between successive ships is 15–145 min. Unloading time per ship varies from 45 to 90 min.

Table 5.16 Harbor system simulation results for 100 ships

Average time of a ship in the harbor	74	62	64	67	67	73
Maximum time of a ship in the harbor	161	116	167	178	173	190
Average waiting time of a ship	19	6	10	12	12	16
Maximum waiting time of a ship	102	58	102	110	104	131
Percentage of time dock facilities are idle	0.25	0.33	0.32	0.30	0.31	0.27

Suppose now that we are not satisfied with the assumption that the arrival time between ships (i.e., their interarrival times) and the unloading time per ship are uniformly distributed over the time intervals $15 \leq \text{between}_i \leq 145$ and $45 \leq \text{unload}_i \leq 90$, respectively. We decide to collect experimental data for the harbor system and incorporate the results into our model, as discussed for the demand submodel in the previous section. We observe (hypothetically) 1200 ships using the harbor to unload their cargoes, and we collect the data displayed in Table 5.18.

Following the procedures outlined in Section 5.4, we consecutively add together the probabilities of each individual time interval between arrivals as well as probabilities of each individual unloading time interval. These computations result in the cumulative histograms depicted in Figure 5.10.

Next we use random numbers uniformly distributed over the interval $0 \leq x \leq 1$ to duplicate the various interarrival times and unloading times based on the cumulative histograms. We then use the midpoints of each interval and construct linear splines through adjacent data points. (We ask you to complete this construction in Problem 1.) Because it is easy to calculate the inverse splines directly, we do so and summarize the results in Tables 5.19 and 5.20.

Table 5.17 Harbor system simulation results for 100 ships

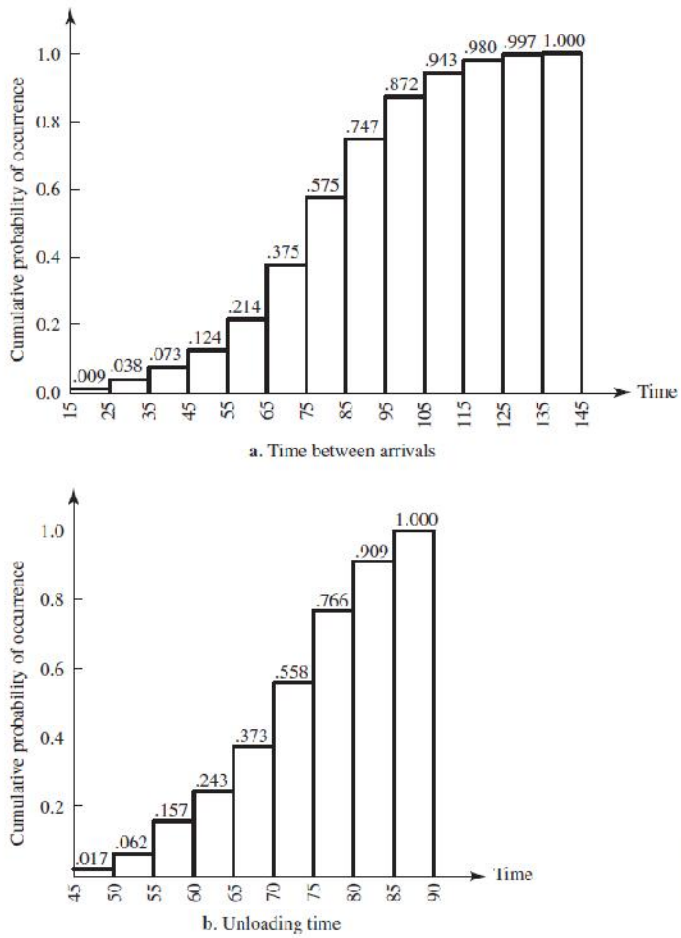
Average time of a ship in the harbor	114	79	96	88	126	115
Maximum time of a ship in the harbor	248	224	205	171	371	223
Average waiting time of a ship	57	24	41	35	71	61
Maximum waiting time of a ship	175	152	155	122	309	173
Percentage of time dock facilities are idle	0.15	0.19	0.12	0.14	0.17	0.06

Note: All times are given in minutes. Time between successive ships is 10–120 min. Unloading time per ship varies from 35 to 75 min.

Table 5.18 Data collected for 1200 ships using the harbor facilities

Time between arrivals	Number of occurrences	Probability of occurrence	Unloading time	Number of occurrences	Probability of occurrence
15 24	11	0.009			
25–34	35	0.029			
35–44	42	0.035	45–49	20	0.017
45–54	61	0.051	50–54	54	0.045
55–64	108	0.090	55–59	114	0.095
65–74	193	0.161	60–64	103	0.086
75–84	240	0.200	65–69	156	0.130
85–94	207	0.172	70–74	223	0.185
95–104	150	0.125	75–79	250	0.208
105–114	85	0.071	80–84	171	0.143
115–124	44	0.037	85–90	109	0.091
125–134	21	0.017		1200	1.000
135–145	3	0.003			
	1200	1.000			

Figure 5.10
Cumulative histograms of the time between ship arrivals and the unloading times, from the data in Table 5.18



Finally, we incorporate our linear spline submodels into the simulation model for the harbor system by generating between_i and unload_i for $i = 1, 2, \ldots, n$ in Steps 1 and 5 of our algorithm, according to the rules displayed in Tables 5.19 and 5.20. Employing these submodels, Table 5.21 gives the results of six independent simulation runs of 100 ships each.

Morning Rush Hour

In the previous example, we initially considered a harbor system with a single facility for unloading ships. Such problems are often called *single-server queues.* In this example, we consider a system with four elevators, illustrating *multiple-server queues.* We discuss the problem and present the algorithm in Appendix B.

Table 5.19 Linear segment submodels provide for the time between arrivals of successive ships as a function of a random number in the interval [0, 1].

Random number interval	Corresponding arrival time	Inverse linear spline
$0 \le x < 0.009$	$15 \le b < 20$	$b = 555.6x + 15.0000$
$0.009 \le x < 0.038$	$20 \le b < 30$	$b = 344.8x + 16.8966$
$0.038 \le x < 0.073$	$30 \le b < 40$	$b = 285.7x + 19.1429$
$0.073 \le x < 0.124$	$40 \le b < 50$	$b = 196.1x + 25.6863$
$0.124 \le x < 0.214$	$50 \le b < 60$	$b = 111.1x + 36.2222$
$0.214 \le x < 0.375$	$60 \le b < 70$	$b = 62.1x + 46.7080$
$0.375 < x < 0.575$	$70 < b < 80$	$b = 50.0x + 51.2500$
$0.575 \le x < 0.747$	$80 \le b < 90$	$b = 58.1x + 46.5698$
$0.747 \le x < 0.872$	$90 \le b < 100$	$b = 80.0x + 30.2400$
$0.872 \le x < 0.943$	$100 \le b < 110$	$b = 140.8x - 22.8169$
$0.943 \le x < 0.980$	$110 \le b < 120$	$b - 270.3x - 144.8649$
$0.980 \le x < 0.997$	$120 \le b < 130$	$b = 588.2x - 456.4706$
$0.997 \le x \le 1.000$	$130 \le b \le 145$	$b = 5000.0x - 4855$

Table 5.20 Linear segment submodels provide for the unloading time of a ship as a function of a random number in the interval [0, 1].

Random number interval	Corresponding unloading time	Inverse linear spline
$0 \le x < 0.017$	$45 \le u < 47.5$	$u = 147x + 45.000$
$0.017 \le x < 0.062$	$47.5 \le u < 52.5$	$u = 111x + 45.611$
$0.062 \le x < 0.157$	$52.5 \le u < 57.5$	$u = 53x + 49.237$
$0.157 \le x < 0.243$	$57.5 \le u < 62.5$	$u = 58x + 48.372$
$0.243 \le x < 0.373$	$62.5 \le u < 67.5$	$u = 38.46x + 53.154$
$0.373 \le x < 0.558$	$67.5 \le u < 72.5$	$u = 27x + 57.419$
$0.558 \le x < 0.766$	$72.5 \le u < 77.5$	$u = 24x + 59.087$
$0.766 \le x < 0.909$	$77.5 \le u < 82.5$	$u = 35x + 50.717$
$0.909 \le x \le 1.000$	$82.5 \le u \le 90$	$u = 82.41x + 7.582$

Consider an office building with 12 floors in a metropolitan area of some city. During e morning rush hour, from 7:50 to 9:10 a.m., workers enter the lobby of the building ıd take an elevator to their floor. There are four elevators servicing the building. The time :tween arrivals of the customers at the building varies in a probabilistic manner every -30 sec, and upon arrival each customer selects the first available elevator (numbered -4). When a person enters an elevator and selects the floor of destination, the elevator aits 15 sec before closing its doors. If another person arrives within the 15-sec interval, e waiting cycle is repeated. If no person arrives within the 15-sec interval, the elevator :parts to deliver all of its passengers. We assume no other passengers are picked up along e way. After delivering its last passenger, the elevator returns to the main floor, picking up) passengers on the way down. The maximum occupancy of an elevator is 12 passengers.

Table 5.21 Harbor system simulation results for 100 ships

Average time of a ship in the harbor	108	95	125	78	123	101
Maximum time of a ship in the harbor	237	188	218	133	250	191
Average waiting time of a ship	38	25	54	9	53	31
Maximum waiting time of a ship	156	118	137	65	167	124
Percentage of time dock facilities are idle	0.09	0.09	0.08	0.12	0.06	0.10

Note: Based on the data exhibited in Table 5.18. All times are given in minutes.

When a person arrives in the lobby and no elevator is available (because all four elevators are transporting their load of passengers), a queue begins to form in the lobby.

The management of the building wants to provide good elevator service to its customers and is interested in exactly what service it is now giving. Some customers claim that they have to wait too long in the lobby before an elevator returns. Others complain that they spend too much time riding the elevator, and still others say that there is considerable congestion in the lobby during the morning rush hour. What is the real situation? Can the management resolve these complaints by a more effective means of scheduling or utilizing the elevators?

We wish to simulate the elevator system using an algorithm for computer implementation that will give answers to the following questions:

1. How many customers are actually being serviced in a typical morning rush hour?
2. If the *waiting time* of a person is the time the person stands in a queue—the time from arrival at the lobby until entry into an available elevator—what are the average and maximum times a person waits in a queue?
3. What is the length of the longest queue? (The answer to this question will provide the management with information about congestion in the lobby.)
4. If the *delivery time* is the time it takes a customer to reach his or her floor after arrival in the lobby, including any waiting time for an available elevator, what are the average and maximum delivery times?
5. What are the average and maximum times a customer actually spends in the elevator?
6. How many stops are made by each elevator? What percentage of the total morning rush hour time is each elevator actually in use?

Linear Programming 1: Geometric Solutions

Consider using the Chebyshev criterion to fit the model $y = cx$ to the following data s

x	1	2	3
y	2	5	8

The optimization problem that determines the parameter c to minimize the largest abso deviation $r_i = |y_i - y(x_i)|$ (residual or error) is the linear program

$$\text{Minimize } r$$

subject to

$$\left.\begin{array}{ll} r-(2-c)\geq 0 & \text{(constraint 1)} \\ r+(2-c)\geq 0 & \text{(constraint 2)} \\ r-(5-2c)\geq 0 & \text{(constraint 3)} \\ r+(5-2c)\geq 0 & \text{(constraint 4)} \\ r-(8-3c)\geq 0 & \text{(constraint 5)} \\ r+(8-3c)\geq 0 & \text{(constraint 6)} \end{array}\right\} \quad ($$

In this section we solve this problem geometrically.

Interpreting a Linear Program Geometrically

Linear programs can include a set of constraints that are linear equations or linear inequalities. Of course, in the case of two decision variables, an equality requires that solutions to the linear program lie precisely on the line representing the equality. What about inequalities? To gain some insight, consider the constraints

$$\begin{aligned} x_1 + 2x_2 &\leq 4 \\ x_1, x_2 &\geq 0 \end{aligned} \tag{7.3}$$

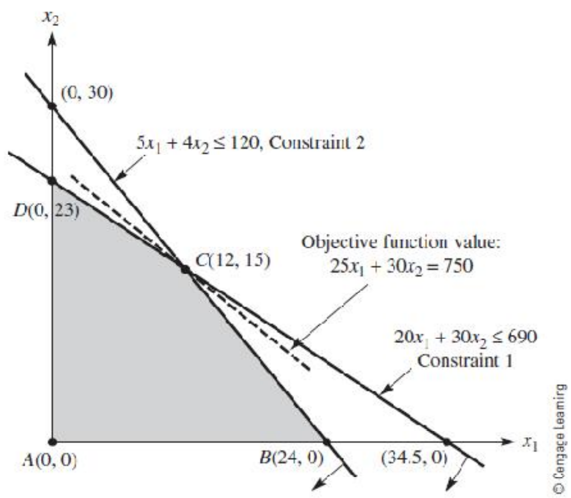
The **nonnegativity** constraints $x_1, x_2 \geq 0$ mean that possible solutions lie in the first quadrant. The inequality $x_1 + 2x_2 \leq 4$ divides the first quadrant into two regions. The *feasible region* is the half-space in which the constraint is satisfied. The feasible region can be found by graphing the equation $x_1 + 2x_2 = 4$ and determining which half-plane is feasible, as shown in Figure 7.2.

If the feasible half-plane fails to be obvious, choose a convenient point (such as the origin) and substitute it into the constraint to determine whether it is satisfied. If it is, then all points on the same side of the line as this point will also satisfy the constraint.

A linear program has the important property that the points satisfying the constraints form a *convex set*. A set is *convex* if for every pair of points in the set, the line segment joining them lies wholly in the set. The set depicted in Figure 7.3a fails to be convex, whereas the set in Figure 7.3b is convex.

An extreme point (corner point) of a convex set is any boundary point in the convex set that is the unique intersection point of two of the (straight-line) boundary segments.

Figure 7.4
The set of points satisfying the constraints of the carpenter's problem form a convex set.



If an optimal solution to a linear program exists, it must occur among the extreme p
of the convex set formed by the set of constraints. The values of the objective function (p
for the carpenter's problem) at the extreme points are

Extreme point	Objective function value
A (0, 0)	\$0
B (24, 0)	600
C (12, 15)	750
D (0, 23)	690

Thus, the carpenter should make 12 tables and 15 bookcases each week to earn a
imum weekly profit of \$750. We provide further geometric evidence later in this se
that extreme point C is optimal.

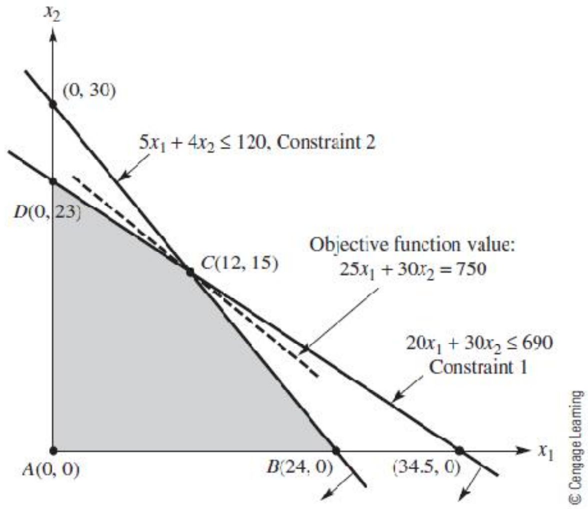
Before considering a second example, let's summarize the ideas presented thu
The constraint set to a linear program is a convex set, which generally contains an in
number of feasible points to the linear program. If an optimal solution to the linear pro
exists, it must be taken on at one or more of the extreme points. Thus, to find an op
solution, we choose from among all the extreme points the one with the best value fo
objective function.

EXAMPLE 2 *A Data-Fitting Problem*

Let's now solve the linear program represented by Equation (7.2). Given the model y =
and the data set

x	1	2	3
y	2	5	8

Figure 7.4
The set of points satisfying the constraints of the carpenter's problem form a convex set.



If an optimal solution to a linear program exists, it must occur among the extreme points of the convex set formed by the set of constraints. The values of the objective function (profit for the carpenter's problem) at the extreme points are

Extreme point	Objective function value
$A\ (0,0)$	\$0
$B\ (24,0)$	600
$C\ (12,15)$	750
$D\ (0,23)$	690

Thus, the carpenter should make 12 tables and 15 bookcases each week to earn a maximum weekly profit of \$750. We provide further geometric evidence later in this section that extreme point C is optimal.

Before considering a second example, let's summarize the ideas presented thus far. The constraint set to a linear program is a convex set, which generally contains an infinite number of feasible points to the linear program. If an optimal solution to the linear program exists, it must be taken on at one or more of the extreme points. Thus, to find an optimal solution, we choose from among all the extreme points the one with the best value for the objective function.

we wish to find a value for c such that the resulting largest absolute deviation r is as small as possible. In Figure 7.5 we graph the set of six constraints

$$
\begin{aligned}
r-(2-c) &\geq 0 && \text{(constraint 1)}\\
r+(2-c) &\geq 0 && \text{(constraint 2)}\\
r-(5-2c) &\geq 0 && \text{(constraint 3)}\\
r+(5-2c) &\geq 0 && \text{(constraint 4)}\\
r-(8-3c) &\geq 0 && \text{(constraint 5)}\\
r+(8-3c) &\geq 0 && \text{(constraint 6)}
\end{aligned}
$$

by first graphing the equations

$$
\begin{aligned}
r-(2-c) &= 0 && \text{(constraint 1 boundary)}\\
r+(2-c) &= 0 && \text{(constraint 2 boundary)}\\
r-(5-2c) &= 0 && \text{(constraint 3 boundary)}\\
r+(5-2c) &= 0 && \text{(constraint 4 boundary)}\\
r-(8-3c) &= 0 && \text{(constraint 5 boundary)}\\
r+(8-3c) &= 0 && \text{(constraint 6 boundary)}
\end{aligned}
$$

We note that constraints 1, 3, and 5 are satisfied above and to the right of the graph of their boundary equations. Similarly, constraints 2, 4, and 6 are satisfied above and to the left of their boundary equations.

The intersection of all the feasible regions for constraints 1–6 forms a convex set in the c, r plane, with extreme points labeled A–C in Figure 7.5. The point A is the intersection of constraint 5 and the r-axis: $r-(8-3c)=0$ and $c=0$, or $A=(0,8)$. Similarly, B is the intersection of constraints 5 and 2:

$$
\begin{aligned}
r-(8-3c) &= 0 \quad \text{or} \quad r+3c=8\\
r+(2-c) &= 0 \quad \text{or} \quad r-c=-2
\end{aligned}
$$

yielding $c = \frac{5}{2}$ and $r = \frac{1}{2}$, or $B = (\frac{5}{2}, \frac{1}{2})$. Finally, C is the intersection of constraints 2 and 4 yielding $C = (3, 1)$. Note that the set is *unbounded*. (We discuss unbounded convex sets later.) If an optimal solution to the problem exists, at least one extreme point must take on the optimal solution. We now evaluate the objective function $f(r) = r$ at each of the three extreme points.

Extreme point	Objective function value
(c, r)	$f(r) = r$
A	8
B	$\frac{1}{2}$
C	1

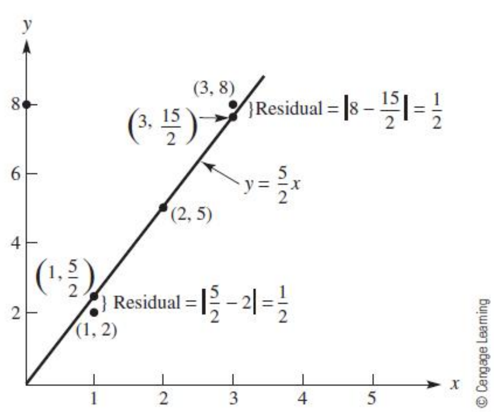
The extreme point with the smallest value of r is the extreme point B with coordinates $(\frac{5}{2}, \frac{1}{2})$. Thus, $c = \frac{5}{2}$ is the optimal value of c. No other value of c will result in a largest absolute deviation as small as $|r_{max}| = \frac{1}{2}$.

Model Interpretation

Let's interpret the optimal solution for the data-fitting problem in Example 2. Resolving the linear program, we obtained a value of $c = \frac{5}{2}$ corresponding to the model $y = \frac{5}{2}x$. Furthermore, the objective function value $r = \frac{1}{2}$ should correspond to the largest deviation resulting from the fit. Let's check to see if that is true.

The data points and the model $y = \frac{5}{2}x$ are plotted in Figure 7.6. Note that a largest deviation of $r_i = \frac{1}{2}$ occurs for both the first and third data points. Fix one end of a ruler at the origin. Now rotate the ruler to convince yourself geometrically that no other line passing through the origin can yield a smaller largest absolute deviation. Thus, the model $y = \frac{5}{2}x$ is optimal by the Chebyshev criterion.

Figure 7.6
The line $y = (5/2)x$ results in a largest absolute deviation $r_{max} = \frac{1}{2}$, the smallest possible r_{max}.



Empty and Unbounded Feasible Regions

We have been careful to say that if an optimal solution to the linear program exists, at least one of the extreme points must take on the optimal value for the objective function. When does an optimal solution fail to exist? Moreover, when does more than one optimal solution exist?

If the feasible region is empty, no feasible solution can exist. For example, given the constraints

$$x_1 \leq 3$$

and

$$x_1 \geq 5$$

there is no value of x_1 that satisfies both of them. We say that such constraint sets are *inconsistent*.

There is another reason an optimal solution may fail to exist. Consider Figure 7.5 and the constraint set for the data-fitting problem in which we noted that the feasible region is *unbounded* (in the sense that either x_1 or x_2 can become arbitrarily large). Then it would be impossible to

$$\text{Maximize } x_1 + x_2$$

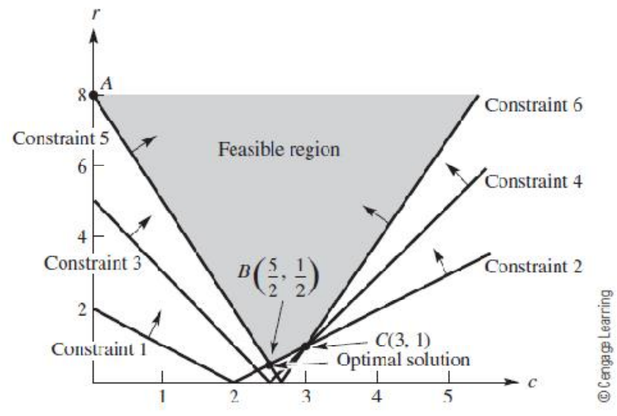
over the feasible region because x_1 and x_2 can take on arbitrarily large values. Note, however, that even though the feasible region is unbounded, an optimal solution *does* exist for the objective function we considered in Example 2, so it is not *necessary* for the feasible region to be bounded for an optimal solution to exist.

EXAMPLE 2 *A Data-Fitting Problem*

Let's now solve the linear program represented by Equation (7.2). Given the model $y = cx$ and the data set

x	1	2	3
y	2	5	8

Figure 7.5
The feasible region for fitting $y = cx$ to a collection of data



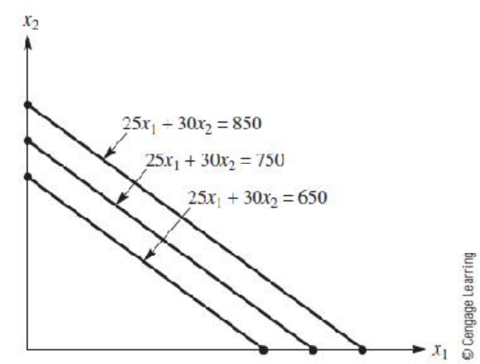
Level Curves of the Objective Function

Consider again the carpenter's problem. The objective function is $25x_1 + 30x_2$ and in Figure 7.7 we plot the lines

$$25x_1 + 30x_2 = 650$$
$$25x_1 + 30x_2 = 750$$
$$25x_1 + 30x_2 = 850$$

in the first quadrant.

Figure 7.7
The level curves of the objective function f are parallel line segments in the first quadrant; the objective function either increases or decreases as we move in a direction perpendicular to the level curves.



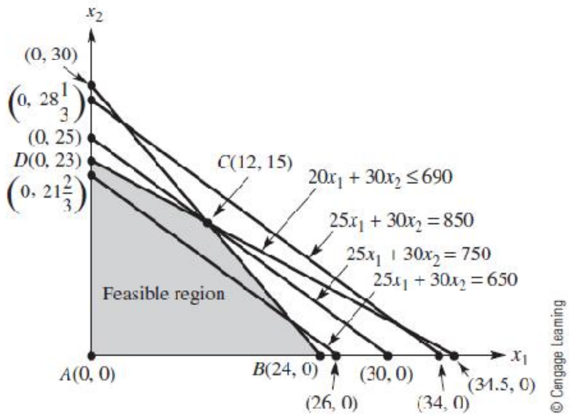
Note that the objective function has constant values along these line segments. The line segments are called **level curves** of the objective function. As we move in a direction perpendicular to these line segments, the objective function either increases or decreases. Now superimpose the constraint set from the carpenter's problem

$$\begin{aligned} 20x_1 + 30x_2 &\leq 690 \quad \text{(lumber)} \\ 5x_1 + 4x_2 &\leq 120 \quad \text{(labor)} \\ x_1, x_2 &\geq 0 \quad \text{(nonnegativity)} \end{aligned}$$

onto these level curves (Figure 7.8). Notice that the level curve with value 750 is the one that intersects the feasible region exactly once at the extreme point $C(12, 15)$.

x_2

Figure 7.8
The level curve $25x_1 + 30x_2 = 750$ is tangent to the feasible region at extreme point C.



Can there be more than one optimal solution? Consider the following slight variatior of the carpenter's problem in which the labor constraint has been changed:

$$\text{Maximize } 25x_1 + 30x_2$$

subject to

$$\begin{aligned} 20x_1 + 30x_2 &\leq 690 \quad &\text{(lumber)} \\ 5x_1 + 6x_2 &\leq 150 \quad &\text{(labor)} \\ x_1, x_2 &\geq 0 \quad &\text{(nonnegativity)} \end{aligned}$$

The constraint set and the level curve $25x_1 + 30x_2 = 750$ are graphed in Figure 7.9. Notice that the level curve and boundary line for the labor constraint coincide. Thus, both extreme points B and C have the same objective function value of 750, which is optimal. In fact the entire line segment BC coincides with the level curve $25x_1 + 30x_2 = 750$. Thus, there are infinitely many optimal solutions to the linear program, all along line segment BC.

In Figure 7.10 we summarize the general two-dimensional case for optimizing a linea function on a convex set. The figure shows a typical convex set together with the level curves of a linear objective function. Figure 7.10 provides geometric intuition for the following fundamental theorem of linear programming.

Figure 7.9

The line segment BC coincides with the level curve $25x_1 + 30x_2 = 750$; every point between extreme points C and B, as well as extreme points C and B, is an optimal solution.

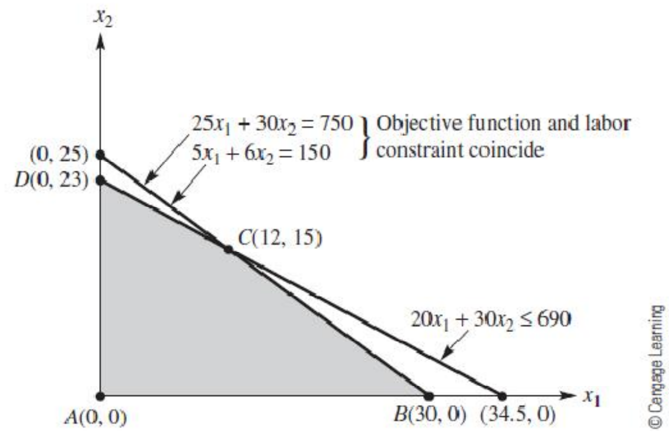
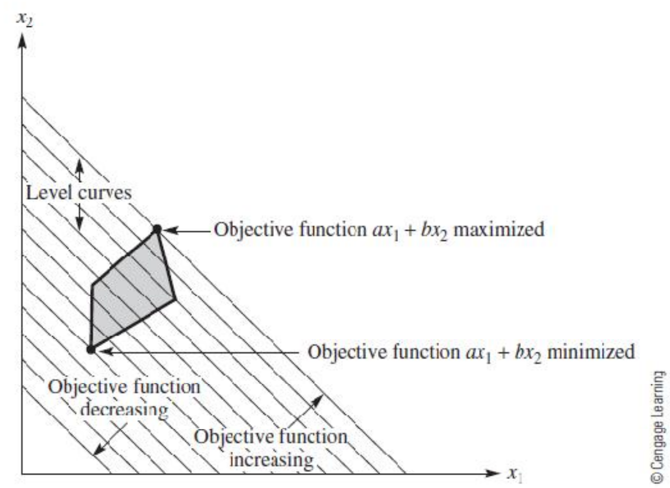


Figure 7.10

A linear function assumes its maximum and minimum values on a nonempty and bounded convex set at an extreme point.



Linear Programming II: Algebraic Solutions

The graphical solution to the carpenter's problem suggests a rudimentary procedure for finding an optimal solution to a linear program with a nonempty and bounded feasible region:

1. Find all intersection points of the constraints.
2. Determine which intersection points, if any, are feasible to obtain the extreme points.
3. Evaluate the objective function at each extreme point.
4. Choose the extreme point(s) with the largest (or smallest) value for the objective function.

To implement this procedure algebraically, we must characterize the intersection points and the extreme points.

The convex set depicted in Figure 7.11 consists of three linear constraints (plus the two nonnegativity constraints). The nonnegative variables y_1, y_2, and y_3 indicated in the figure measure the degree by which a point satisfies the constraints 1, 2, and 3, respectively. The variable y_i is added to the left side of inequality constraint i to convert it to an equality. Thus, $y_2 = 0$ characterizes those points that lie precisely on constraint 2, and a negative value for y_2 indicates the violation of constraint 2. Likewise, the decision variables x_1 and x_2 are constrained to nonnegative values. Thus, the values of the decision variables x_1 and x_2 measure the degree of satisfaction of the nonnegativity constraints, $x_1 \geq 0$ and $x_2 \geq 0$. Note that along the x_1-axis, the decision variable x_2 is 0. Now consider the values for the entire set of variables $\{x_1, x_2, y_1, y_2, y_3\}$. If two of the variables simultaneously have the value 0, then we have characterized an **intersection point** in the x_1x_2-plane. All (possible) intersection points can be determined systematically by setting all possible distinguishable pairs of the five variables to zero and solving for the remaining three dependent variables. If a solution to the resulting system of equations exists, then it must be an intersection point, which may

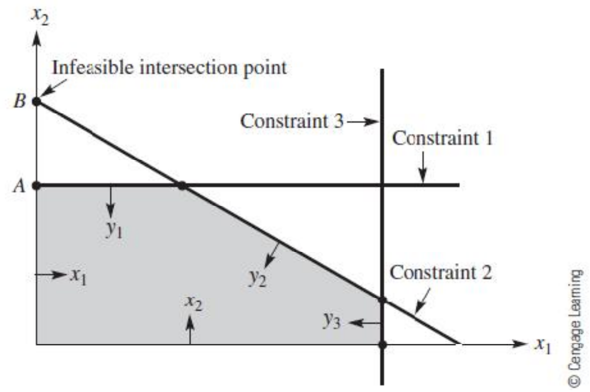


Figure 7.11
The variables x_1, x_2, y_1, y_2, and y_3 measure the satisfaction of each of the constraints; intersection point A is characterized by $y_1 = x_1 = 0$; intersection point B is not feasible because y_1 is negative; the intersection points surrounding the shaded region are all feasible because none of the five variables is negative there.

or may not be a **feasible solution**. A negative value for any of the five variables indicates that a constraint is not satisfied. Such an intersection point would be **infeasible**. For example, the intersection point B, where $y_2 = 0$ and $x_1 = 0$, gives a negative value for y_1 and hence is not feasible. Other pairs of variables such as x_1 and y_3, cannot simultaneously be zero because they represent constraints that are parallel lines. Let's illustrate the procedure by solving the carpenter's problem algebraically.

PLE 1 *Solving the Carpenter's Problem Algebraically*

The carpenter's model is

$$\text{Maximize } 25x_1 + 30x_2$$

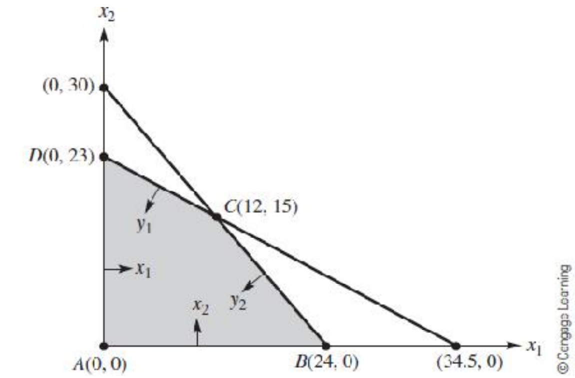
subject to

$$\begin{aligned} 20x_1 + 30x_2 &\leq 690 \quad \text{(lumber)} \\ 5x_1 + 4x_2 &\leq 120 \quad \text{(labor)} \\ x_1, x_2 &\geq 0 \quad \text{(nonnegativity)} \end{aligned}$$

We convert each of the first two inequalities to equations by adding new nonnegative "slack" variables y_1 and y_2. If either y_1 or y_2 is negative, the constraint is not satisfied. Thus, the problem becomes

$$\text{Maximize } 25x_1 + 30x_2$$

Figure 7.12
The variables $\{x_1, x_2, y_1, y_2\}$ measure the satisfaction of each constraint; an intersection point is characterized by setting two of the variables to zero.



subject to

$$
\begin{aligned}
20x_1 + 30x_2 + y_1 &= 690 \\
5x_1 + 4x_2 + y_2 &= 120 \\
x_1, x_2, y_1, y_2 &\geq 0
\end{aligned}
$$

We now consider the entire set of four variables $\{x_1, x_2, y_1, y_2\}$, which are interpreted geometrically in Figure 7.12. To determine a possible intersection point in the x_1x_2-plane, assign two of the four variables the value zero. There are $\frac{4!}{2!\,2!} = 6$ possible intersection points to consider in this way (four variables taken two at a time). Let's begin by assigning the variables x_1 and x_2 the value zero, resulting in the following set of equations:

$$
\begin{aligned}
y_1 &= 690 \\
y_2 &= 120
\end{aligned}
$$

which is a feasible intersection point $A(0, 0)$ because all four variables are nonnegative.

For the second intersection point we choose the variables x_1 and y_1 and set them to zero, resulting in the system

$$
\begin{aligned}
30x_2 &= 690 \\
4x_2 + y_2 &= 120
\end{aligned}
$$

that has solution $x_2 = 23$ and $y_2 = 28$, which is also a feasible intersection point $D(0, 23)$.

For the third intersection point we choose x_1 and y_2 and set them to zero, yielding the system

$$
\begin{aligned}
30x_2 + y_1 &= 690 \\
4x_2 &= 120
\end{aligned}
$$

with solution $x_2 = 30$ and $y_1 = -210$. Thus, the first constraint is violated by 210 units, indicating that the intersection point $(0, 30)$ is infeasible.

In a similar manner, choosing y_1 and y_2 and setting them to zero gives $x_1 = 12$ and $x_2 = 15$, corresponding to the intersection point $C(12, 15)$, which is feasible. Our fifth choice is to choose the variables x_2 and y_1 and set them to zero, giving values of $x_1 = 34.5$ and $y_2 = -52.5$, so the second constraint is not satisfied. Thus, the intersection point $(34.5, 0)$ is infeasible.

Finally we determine the sixth intersection point by setting the variables x_2 and y_2 to zero to determine $x_1 = 24$ and $y_1 = 210$; therefore, the intersection point $B(24, 0)$ is feasible.

In summary, of the six possible intersection points in the x_1x_2-plane, four were found to be feasible. For the four we find the value of the objective function by substitution:

Extreme point	Value of objective function
$A(0, 0)$	\$0
$D(0, 23)$	690
$C(12, 15)$	750
$B(24, 0)$	600

Our procedure determines that the optimal solution to maximize the profit is $x_1 = 12$ and $x_2 = 15$. That is, the carpenter should make 12 tables and 15 bookcases for a maximum profit of \$750.

Linear Programming III: The Simplex Method

So far we have learned to find an optimal extreme point by searching among all possible intersection points associated with the decision and slack variables. Can we reduce the number of intersection points we actually consider in our search? Certainly, once we find an initial feasible intersection point, we need not consider a potential intersection point that fails to improve the value of the objective function. Can we test the optimality of our current solution against other possible intersection points? Even if an intersection point promises to be more optimal than the current extreme point, it is of no interest if it violates one or more of the constraints. Is there a test to determine whether a proposed intersection point is feasible? The **Simplex Method**, developed by George Dantzig, incorporates both *optimality* and *feasibility* tests to find the optimal solution(s) to a linear program (if one exists).

An **optimality test** shows whether or not an intersection point corresponds to a value of the objective function better than the best value found so far.

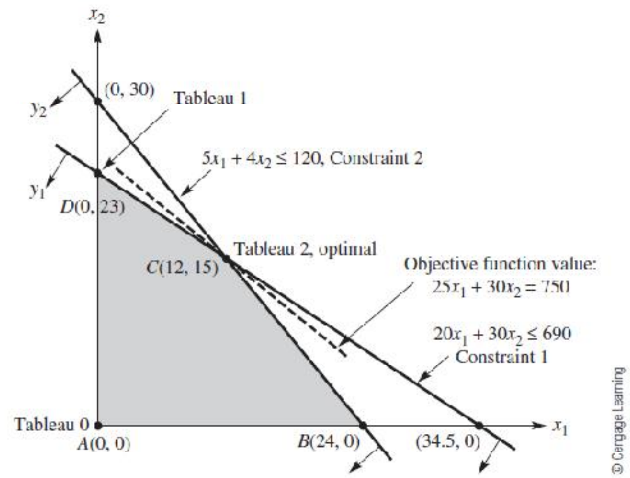
A **feasibility test** determines whether the proposed intersection point is feasible.

To implement the Simplex Method we first separate the decision and slack variables into two nonoverlapping sets that we call the **independent** and **dependent** sets. For the particular linear programs we consider, the original independent set will consist of the decision variables, and the slack variables will belong to the dependent set.

Steps of the Simplex Method

1. **Tableau Format:** Place the linear program in Tableau Format, as explained later.
2. **Initial Extreme Point:** The Simplex Method begins with a known extreme point, usually the origin $(0, 0)$.
3. **Optimality Test:** Determine whether an adjacent intersection point improves the value of the objective function. If not, the current extreme point is optimal. If an improvement is possible, the optimality test determines which variable currently in the independent set (having value zero) should *enter* the dependent set and become nonzero.
4. **Feasibility Test:** To find a new intersection point, one of the variables in the dependent set must *exit* to allow the entering variable from Step 3 to become dependent. The feasibility test determines which current dependent variable to choose for exiting, ensuring feasibility.
5. **Pivot:** Form a new, equivalent system of equations by eliminating the new dependent variable from the equations do not contain the variable that exited in Step 4. Then set the new independent variables to zero in the new system to find the values of the new dependent variables, thereby determining an intersection point.
6. **Repeat Steps 3–5** until an optimal extreme point is found.

Figure 7.13
The set of points satisfying the constraints of a linear program (the shaded region) form a convex set.



Before detailing each of the preceding steps, let's examine the carpenter's problem (Figure 7.13). The origin is an extreme point, so we choose it as our starting point. Thus, x_1 and x_2 are the current arbitrary independent variables and are assigned the value zero, whereas y_1 and y_2 are the current dependent variables with values of 690 and 120, respectively. The optimality test determines whether a current independent variable assigned the value zero could improve the value of the objective function if it is made dependent and positive. For example, either x_1 or x_2, if made positive, would improve the objective function value. (They have positive coefficients in the objective function we are trying to maximize.) Thus, the optimality test determines a promising variable to enter the dependent set. Later, we give a rule of thumb for choosing which independent variable to enter when more than one candidate exists. In the carpenter's problem at hand, we select x_2 as the new dependent variable.

The variable chosen for entry into the dependent set by the optimality condition replaces one of the current dependent variables. The feasibility condition determines which exiting variable this entering variable replaces. Basically, the entering variable replaces whichever current dependent variable can assume a zero value while maintaining nonnegative values for all the remaining dependent variables. That is, the feasibility condition ensures that the new intersection point will be feasible and hence an extreme point. In Figure 7.13, the feasibility test would lead us to the intersection point $(0, 23)$, which is feasible, and not to $(0, 30)$, which is infeasible. Thus, x_2 replaces y_1 as a dependent or nonzero variable. Therefore, x_2 enters and y_1 exits the set of dependent variables.

Computational Efficiency

The feasibility test does not require actual computation of the values of the dependent variables when selecting an exiting variable for replacement. Instead, we will see that an appropriate exiting variable is selected by quickly determining whether any variable becomes negative if the dependent variable being considered for replacement is assigned the value zero (a ratio test that will be explained later). If any variable would become negative, then the dependent variable under consideration cannot be replaced by the entering variable if feasibility is to be maintained. Once a set of dependent variables corresponding to a more optimal extreme point is found from the optimality and feasibility tests, the values of the new dependent variables are determined by pivoting. The pivoting process essentially solves an equivalent system of equations for the new dependent variables after the exchange of the entering and exiting variables in the dependent set. The values of the new dependent variables are obtained by assigning the independent variables the value zero. Note that only one dependent variable is replaced at each stage. *Geometrically, the Simplex Method proceeds from an initial extreme point to an adjacent extreme point until no adjacent extreme point is more optimal.* At that time, the current extreme point is an optimal solution. We now detail the steps of the Simplex Method.

STEP 1 **TABLEAU FORMAT** Many formats exist for implementing the Simplex Method. The format we use assumes that the objective function is to be maximized and that the constraints are less than or equal to inequalities. (If the problem is not expressed initially in this format, it can easily be changed to this format.) For the carpenter's example, the problem is

$$\text{Maximize } 25x_1 + 30x_2$$

subject to

$$\begin{aligned} 20x_1 + 30x_2 &\le 690 \\ 5x_1 + 4x_2 &\le 120 \\ x_1, x_2 &\ge 0 \end{aligned}$$

Next we adjoin a new constraint to ensure that any solution improves the best value of the objective function found so far. Take the initial extreme point as the origin, where the value of the objective function is zero. We want to constrain the objective function to be better than its current value, so we require

$$25x_1 + 30x_2 \ge 0$$

Because all the constraints must be $\le$ inequalities, multiply the new constraint by -1 and adjoin it to the original constraint set:

$$\begin{aligned} 20x_1 + 30x_2 &\le 690 \quad &&\text{(constraint 1, lumber)} \\ 5x_1 + 4x_2 &\le 120 \quad &&\text{(constraint 2, labor)} \\ -25x_1 - 30x_2 &\le 0 \quad &&\text{(objective function constraint)} \end{aligned}$$

The Simplex Method implicitly assumes that all variables are nonnegative, so we do not repeat the nonnegativity constraints in the remainder of the presentation.

Next, we convert each inequality to an equality by adding a *nonnegative* new variable y_i (or z), called a *slack variable* because it measures the slack or degree of satisfaction of the constraint. A negative value for y_i indicates the constraint is not satisfied. (We use the variable z for the objective function constraint to avoid confusion with the other constraints.)

This process gives the *augmented constraint set*

$$\begin{aligned} 20x_1 + 30x_2 + y_1 &= 690 \\ 5x_1 + 4x_2 + y_2 &= 120 \\ -25x_1 - 30x_2 + z &= 0 \end{aligned}$$

where the variables x_1, x_2, y_1, y_2 are nonnegative. The value of the variable z represents the value of the objective function, as we shall see later. (Note from the last equation that $z = 25x_1 + 30x_2$ is the value of the objective function.)

STEP 2 **INITIAL EXTREME POINT** Because there are two decision variables, all possible intersection points lie in the x_1x_2-plane and can be determined by setting two of the variables $\{x_1, x_2, y_1, y_2\}$ to zero. (The variable z is *always* a dependent variable and represents the value of the objective function at the extreme point in question.) The origin is feasible and corresponds to the extreme point characterized by $x_1 = x_2 = 0$, $y_1 = 690$, and $y_2 = 120$. Thus, x_1 and x_2 are independent variables assigned the value 0; y_1, y_2, and z are dependent variables whose values are then determined. As we will see, z conveniently records the current value of the objective function at the extreme points of the convex set in the x_1x_2-plane as we compute them by elimination.

STEP 3 **THE OPTIMALITY TEST FOR CHOOSING AN ENTERING VARIABLE** In the preceding format, a negative coefficient in the last (or objective function) equation indicates that the corresponding variable could improve the current objective function value. Thus, the coefficients -25 and -30 indicate that either x_1 or x_2 could enter and improve the current objective function value of $z = 0$. (The current constraint corresponds to $z = 25x_1 + 30x_2 \geq 0$, with x_1 and x_2 currently independent and 0.) When more than one candidate exists for the entering variable, a rule of thumb for selecting the variable to enter the dependent set is to select that variable with the largest (in absolute value) negative coefficient in the objective function row. If no negative coefficients exist, the current solution is optimal. In the case at hand, we choose x_2 as the new entering variable. That is, x_2 will increase from its current value of zero. The next step determines how great an increase is possible.

STEP 4 **THE FEASIBILITY CONDITION FOR CHOOSING AN EXITING VARIABLE** The entering variable x_2 (in our example) must replace either y_1 or y_2 as a dependent variable (because z *always* remains the third dependent variable). To determine which of these variables is to exit the dependent set, first divide the right-hand-side values 690 and 120 (associated with the original constraint inequalities) by the components for the entering variable in each inequality (30 and 4, respectively, in our example) to obtain the ratios $\frac{690}{30} = 23$ and $\frac{120}{4} = 30$. From the subset of ratios that are positive (both in this case), the variable corresponding to the minimum ratio is chosen for replacement (y_1, which corresponds to 23 in this case). *The ratios represent the value the entering variable would obtain if the corresponding exiting variable were assigned the value* 0. Thus, only positive values are considered and the smallest positive value is chosen so as not to drive any variable negative. For instance, if y_2 were chosen as the exiting variable and assigned the value 0, then x_2 would assume a value 30 as the new dependent variable. However, then y_1 would be negative, indicating that the intersection point $(0, 30)$ does not satisfy the first constraint. Note that the intersection point $(0, 30)$ is not feasible in Figure 7.13. The **minimum positive ratio rule** illustrated previously obviates enumeration of any infeasible intersection points. In the case at hand, the dependent variable corresponding to the smallest ratio 23 is y_1, so it becomes the exiting variable. Thus, x_2, y_2, and z form the new set of dependent variables, and x_1 and y_1 form the new set of independent variables.

STEP 5 **PIVOTING TO SOLVE FOR THE NEW DEPENDENT VARIABLE VALUES** Next we derive a new (equivalent) system of equations by eliminating the entering variable x_2 in all the equations of the previous system that do not contain the exiting variable y_1. There are numerous ways to execute this step, such as the method of elimination used in Section 7.3. Then we find the values of the dependent variables x_2, y_2, and z when the independent variables x_1 and y_1 are assigned the value 0 in the new system of equations. This is called the **pivoting procedure**. The values of x_1 and x_2 give the new extreme point (x_1, x_2), and z is the (improved) value of the objective function at that point.

After performing the pivot, apply the optimality test again to determine whether another candidate entering variable exists. If so, choose an appropriate one and apply the feasibility test to choose an exiting variable. Then the pivoting procedure is performed again. The process is repeated until no variable has a negative coefficient in the objective function row. We now summarize the procedure and use it to solve the carpenter's problem.

Summary of the Simplex Method

STEP 1 **PLACE THE PROBLEM IN TABLEAU FORMAT.** Adjoin slack variables as needed to convert inequality constraints to equalities. Remember that all variables are nonnegative. Include the objective function constraint as the last constraint, including its slack variable z.

STEP 2 **FIND ONE INITIAL EXTREME POINT.** (For the problems we consider, the origin will be an extreme point.)

STEP 3 **APPLY THE OPTIMALITY TEST.** Examine the last equation (which corresponds to the objective function). If all its coefficients are nonnegative, then stop: The current extreme point is optimal. Otherwise, some variables have negative coefficients, so choose the variable with the largest (in absolute value) negative coefficient as the new entering variable.

STEP 4 **APPLY THE FEASIBILITY TEST.** Divide the current right-hand-side values by the corresponding coefficient values of the entering variable in each equation. Choose the exiting variable to be the one corresponding to the smallest positive ratio after this division.

STEP 5 **PIVOT.** Eliminate the entering variable from all the equations that do not contain the exiting variable. (For example, you can use the elimination procedure presented in Section 7.2.) Then assign the value 0 to the variables in the new independent set (consisting of the exited variable and the variables remaining after the entering variable has left to become dependent). The resulting values give the new extreme point (x_1, x_2) and the objective function value z for that point.

STEP 6 **REPEAT STEPS 3–5** until an optimal extreme point is found.

EXAMPLE 1 *The Carpenter's Problem Revisited*

STEP 1 The Tableau Format gives

$$\begin{aligned} 20x_1 + 30x_2 + y_1 &= 690 \\ 5x_1 + 4x_2 + y_2 &= 120 \\ -25x_1 - 30x_2 + z &= 0 \end{aligned}$$

STEP 2 The origin $(0, 0)$ is an initial extreme point for which the independent variables are x_1 = $x_2 = 0$ and the dependent variables are $y_1 = 690$, $y_2 = 120$, and $z = 0$.

STEP 3 We apply the optimality test to choose x_2 as the variable entering the dependent set becau it corresponds to the negative coefficient with the largest absolute value.

STEP 4 Applying the feasibility test, we divide the right-hand-side values 690 and 120 by th components for the entering variable x_2 in each equation (30 and 4, respectively), yieldir the ratios $\frac{690}{30} = 23$ and $\frac{120}{4} = 30$. The smallest positive ratio is 23, corresponding to th first equation that has the slack variable y_1. Thus, we choose y_1 as the exiting depende variable.

STEP 5 We pivot to find the values of the new dependent variables x_2, y_2, and z when the independe variables x_1 and y_1 are set to the value 0. After eliminating the new dependent variable x from each previous equation that does not contain the exiting variable y_1, we obtain th equivalent system

$$
\begin{aligned}
\frac{2}{3}x_1 + x_2 + \frac{1}{30}y_1 &= 23 \\
\frac{7}{3}x_1 - \frac{2}{15}y_1 + y_2 &= 28 \\
-5x_1 + y_1 + z &= 690
\end{aligned}
$$

Setting $x_1 = y_1 = 0$, we determine $x_2 = 23$, $y_2 = 28$, and $z = 690$. These results give th extreme point $(0, 23)$ where the value of the objective function is $z = 690$.

Applying the optimality test again, we see that the current extreme point $(0, 23)$ is n optimal (because there is a negative coefficient -5 in the last equation corresponding to th variable x_1). Before continuing, observe that we really do not need to write out the enti symbolism of the equations in each step. We merely need to know the coefficient values a sociated with the variables in each of the equations together with the right-hand side. A tab format, or *tableau*, is commonly used to record these numbers. We illustrate the completic of the carpenter's problem using this format, where the headers of each column designate th variables; the abbreviation RHS heads the column where the values of the right-hand sic appear. We begin with Tableau 0, corresponding to the initial extreme point at the origin.

Tableau 0 (Original Tableau)

x_1	x_2	y_1	y_2	z	RHS
20	30	1	0	0	$690 (= y_1)$
5	4	0	1	0	$120 (= y_2)$
-25	-30	0	0	1	$0 (= z)$

Dependent variables: $\{y_1, y_2, z\}$
Independent variables: $x_1 = x_2 = 0$
Extreme point: $(x_1, x_2) = (0, 0)$
Value of objective function: $z = 0$

Optimality Test The entering variable is x_2 (corresponding to -30 in the last row).

Feasibility Test Compute the ratios for the RHS divided by the coefficients in the colum labeled x_2 to determine the minimum positive ratio.

Entering variable

x_1	x_2	y_1	y_2	z	RHS	Ratio
20	30	1	0	0	690	23 (= 690/30) ← Exiting variable
5	4	0	1	0	120	30 (= 120/4)
−25	−30	0	0	1	0	*

Choose y_1 corresponding to the minimum positive ratio 23 as the exiting variable.

Pivot Divide the row containing the exiting variable (the first row in this case) by the coefficient of the entering variable in that row (the coefficient of x_2 in this case), giving a coefficient of 1 for the entering variable in this row. Then eliminate the entering variable x_2 from the remaining rows (which do not contain the exiting variable y_1 and have a zero coefficient for it). The results are summarized in the next tableau, where we use five-place decimal approximations for the numerical values.

Tableau 1

x_1	x_2	y_1	y_2	z	RHS
0.66667	1	0.03333	0	0	23 $(= x_2)$
2.33333	0	−0.13333	1	0	28 $(= y_2)$
−5.00000	0	1.00000	0	1	690 $(= z)$

Dependent variables: $\{x_2, y_2, z\}$
Independent variables: $x_1 = y_1 = 0$
Extreme point: $(x_1, x_2) = (0, 23)$
Value of objective function: $z = 690$

The pivot determines that the new dependent variables have the values $x_2 = 23$, $y_2 = 28$, and $z = 690$.

Optimality Test The entering variable is x_1 (corresponding to the coefficient −5 in the last row).

Feasibility Test Compute the ratios for the RHS.

Entering variable

x_1	x_2	y_1	y_2	z	RHS	Ratio
0.66667	1	0.03333	0	0	23	34.5 (= 23/0.66667)
2.33333	0	−0.13333	1	0	28	12.0 (= 28/2.33333) ← Exiting variable
−5.00000	0	1.00000	0	1	690	*

Choose y_2 as the exiting variable because it corresponds to the minimum positive ratio 12.

Pivot Divide the row containing the exiting variable (the second row in this case) by the coefficient of the entering variable in that row (the coefficient of x_1 in this case), giving a coefficient of 1 for the entering variable in this row. Then eliminate the entering variable x_1 from the remaining rows (which do not contain the exiting variable y_2 and have a zero coefficient for it). The results are summarized in the next tableau.

Tableau 2

x_1	x_2	y_1	y_2	z	RHS
0	1	0.071429	−0.28571	0	15 $(= x_2)$
1	0	−0.057143	0.42857	0	12 $(= x_1)$
0	0	0.714286	2.14286	1	750 $(= z)$

Dependent variables: $\{x_2, x_1, z\}$
Independent variables: $y_1 = y_2 = 0$
Extreme point: $(x_1, x_2) = (12, 15)$
Value of objective function: $z = 750$

Optimality Test Because there are no negative coefficients in the bottom row, $x_1 = 12$ and $x_2 = 15$ gives the optimal solution $z = \$750$ for the objective function. Note that starting with an initial extreme point, we had to enumerate only two of the possible six intersection points. The power of the Simplex Method is its reduction of the computations required to find an optimal extreme point.

Questions	opt1	opt2	opt3	opt4	opt5	opt6	Answer
The model simulates the performance of the harbor with time using ____ main models	2	3	4	6			2
First model of the harbor system is track of the ships____	stable situation	waiting time	position	position and waiting time			position and waiting time
Second model of the Harbor system is calculate the value of the ____	area	volume	action	climatic agents			climatic agents
Time series simulation is also connected with ____ agents	random	unique	climatic	neither climatic or random			climatic
The use of exponential distribution for modeling inter arrival time and widely used in __	bibliography	cryptography	decision	verification			bibliography
Morning rush hour model results are divided on ____ and ____ of the design	verification and optimisation	verification and decision	verification and observation	decision and optimisation			verification and optimisation
In basic model various components of the vector x are called ___ variable	negative	negligible	non negligible	decision			negative
In basic model $f_i(x)$ are called ____ function	objective	negative	decision	non negative			decision
In basic model side condition are typically called ____	constrains	decision variables	non negative restrictions	negative restrictions			constrains
Optimisation problem is said to be linear program if there is ___ objective function	many	more than one	unique	no			no
Optimisation problem is also said to be	unconstrained programming	linear programming	unconstrained or linear programming	either unconstrained nor linear			unconstrained programming
In a linear programming problem has the coefficient of the decision variable in the objective function and each constrains are	unique	un uniform	standard	constant			constant
In linear program optimisation problem satisfies the decision variable are permitted to assume fraction as well as ____ values	constant	complex	real	integer			integer

Questions	opt1	opt2	opt3	opt4	opt5	opt6	Answer
Problems with more than one objective function are called ____ programs	multi objective	either goal or multi objective	multi objective and goal	goal			multi objective
Any optimisation problems that fails to satisfies either one of the multi objective is said to be	ordered	un ordered	linear	non linear			non linear
Time dependent problem in a certain classes are called ___ programmes	stochastic	linear	integer	dynamic			linear
The resulting problem is called an ____ programme	integer	stochastic	linear	real			integer
If the coefficient are not constant but instead of probabilistic in nature , the problem is classified as a ____ program	stochastic	integer	dynamic	linear			stochastic
Multi objective programs are also called as_____ programs	goal	linear	integer	mixed			goal
Integer optimisation programs may restrict ---_______ of decision variable to integer values	one or more	less than one	one	multiple			one or more
_____ method does not solve integer or mixed integer problem directly	simplex	binomial	integer	complex			simplex
_____ program is called as an integer programming	resulting	linear	non linear	integer			resulting
________dependent problems in a certain class are called dynamic programs	time	climate	weight	positions			time
Integer programming are also called as ____ program	mixed integer	integer	neither mixed integer or	non linear			mixed integer

Hooke's Law

The magnitude of the force needed to produce a certain elongation of a spring is directly proportional to the amount of this elongation, provided this elongation is not too great. In mathematical form,

$$|F| = ks,$$

where F is the magnitude of the force, s is the amount of elongation, and k is a constant of proportionality which we shall call the *spring constant*.

The spring constant k depends upon the spring under consideration and is a measure of its stiffness. For example, if a 30-lb weight stretches a spring 2 ft, then Hooke's law gives $30 = (k)(2)$; thus for this spring $k = 15$ lb/ft.

When a mass is hung upon a spring of spring constant k and thus produces an elongation of amount s, the force F of the mass upon the spring therefore has magnitude ks. The spring at the same time exerts a force upon the mass called the *restoring force* of the spring. This force is equal in magnitude but opposite in sign to F and hence has magnitude $-ks$.

Let us formulate the problem systematically. Let the coil spring have natural (unstretched) length L. The mass m is attached to its lower end and comes to rest in its equilibrium position, thereby stretching the spring an amount l so that its stretched length is $L + l$. We choose the axis along the line of the spring, with the origin O at the equilibrium position and the positive direction downward. Thus, letting x denote the displacement of the mass from O along this line, we see that x is positive, zero, or negative according to whether the mass is below, at, or above its equilibrium position. (See Figure 5.1.)

Forces Acting Upon the Mass

We now enumerate the various forces that act upon the mass. Forces tending to pull the mass downward are positive, while those tending to pull it upward are negative. The forces are:

1. F_1, the *force of gravity*, of magnitude mg, where g is the acceleration due to gravity. Since this acts in the downward direction, it is positive, and so

$$F_1 = mg. \tag{5.1}$$

2. F_2, the *restoring force* of the spring. Since $x + l$ is the total amount of elongation, by Hooke's law the magnitude of this force is $k(x + l)$. When the mass is *below* the end of the unstretched spring, this force acts in the *upward* direction and so is *negative*. Also,

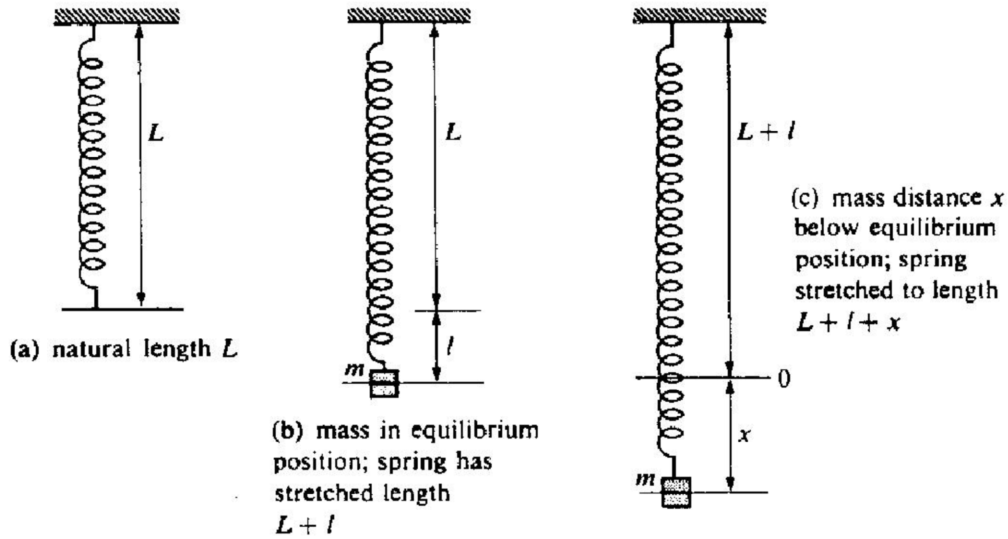


Figure 5.1

for the mass in such a position, $x + l$ is *positive*. Thus, when the mass is *below* the end of the unstretched spring, the restoring force is given by

$$F_2 = -k(x + l). \tag{5.2}$$

This also gives the restoring force when the mass is *above* the end of the unstretched spring, as one can see by replacing each italicized word in the three preceding sentences by its opposite. When the mass is at rest in its equilibrium position the restoring force F_2 is equal in magnitude but opposite in direction to the force of gravity and so is given by $-mg$. Since in this position $x = 0$, Equation (5.2) gives

$$-mg = -k(0 + l)$$

or

$$mg = kl.$$

Replacing kl by mg in Equation (5.2) we see that the restoring force can thus be written as

$$F_2 = -kx - mg. \tag{5.3}$$

3. F_3, the *resisting force* of the medium, called the *damping force*. Although the magnitude of this force is not known *exactly*, it is known that for small velocities it is *approximately* proportional to the magnitude of the velocity:

$$|F_3| = a\left|\frac{dx}{dt}\right|, \tag{5.4}$$

where $a > 0$ is called the *damping constant*. When the mass is moving *downward*, F_3 acts in the *upward* direction (opposite to that of the motion) and so $F_3 < 0$. Also, since m is moving *downward*, x is *increasing* and dx/dt is *positive*. Thus, assuming Equation (5.4) to hold, when the mass is moving *downward*, the damping force is given by

This also gives the damping force when the mass is moving *upward*, as one may see by replacing each italicized word in the three preceding sentences by its opposite.

4. F_4, any *external impressed forces* that act upon the mass. Let us denote the resultant of all such external forces at time t simply by $F(t)$ and write

$$F_4 = F(t). \tag{5.6}$$

We now apply Newton's second law, $F = ma$, where $F = F_1 + F_2 + F_3 + F_4$. Using (5.1), (5.3), (5.5), and (5.6), we find

$$m\frac{d^2x}{dt^2} = mg - kx - mg - a\frac{dx}{dt} + F(t)$$

or

$$m\frac{d^2x}{dt^2} + a\frac{dx}{dt} + kx = F(t). \tag{5.7}$$

This we take as the differential equation for the motion of the mass on the spring. Observe that it is a nonhomogeneous second-order linear differential equation with constant coefficients. If $a = 0$ the motion is called *undamped*; otherwise it is called *damped*. If there are no external impressed forces, $F(t) = 0$ for all t and the motion is called *free*; otherwise it is called *forced*. In the following sections we consider the solution of (5.7) in each of these cases.

5.2 FREE, UNDAMPED MOTION

We now consider the special case of *free, undamped motion*, that is, the case in which both $a = 0$ and $F(t) = 0$ for all t. The differential equation (5.7) then reduces to

$$m\frac{d^2x}{dt^2} + kx = 0, \tag{5.8}$$

where $m(>0)$ is the mass and $k(>0)$ is the spring constant. Dividing through by m and letting $k/m = \lambda^2$, we write (5.8) in the form

$$\frac{d^2x}{dt^2} + \lambda^2 x = 0. \tag{5.9}$$

The auxiliary equation

$$r^2 + \lambda^2 = 0$$

has roots $r = \pm\lambda i$ and hence the general solution of (5.8) can be written

$$x = c_1 \sin \lambda t + c_2 \cos \lambda t, \tag{5.10}$$

where c_1 and c_2 are arbitrary constants.

Let us now assume that the mass was initially displaced a distance x_0 from its equilibrium position and released from that point with initial velocity v_0. Then, in addition to the differential equation (5.8) [or (5.9)], we have the initial conditions

$$x(0) = x_0, \tag{5.11}$$

$$x'(0) = v_0. \tag{5.12}$$

Differentiating (5.10) with respect to t, we have

$$\frac{dx}{dt} = c_1 \lambda \cos \lambda t - c_2 \lambda \sin \lambda t. \tag{5.13}$$

Applying conditions (5.11) and (5.12) to Equations (5.10) and (5.13), respectively, we see at once that

$$c_2 = x_0,$$
$$c_1 \lambda = v_0.$$

Substituting the values of c_1 and c_2 so determined into Equation (5.10) gives the particular solution of the differential equation (5.8) satisfying the conditions (5.11) and (5.12) in the form

$$x = \frac{v_0}{\lambda} \sin \lambda t + x_0 \cos \lambda t.$$

We put this in an alternative form by first writing it as

$$x = c\left[\frac{(v_0/\lambda)}{c} \sin \lambda t + \frac{x_0}{c} \cos \lambda t\right], \tag{5.14}$$

where

$$c = \sqrt{\left(\frac{v_0}{\lambda}\right)^2 + x_0^2} > 0. \tag{5.15}$$

Then, letting

$$\begin{aligned} \frac{(v_0/\lambda)}{c} &= -\sin \phi, \\ \frac{x_0}{c} &= \cos \phi, \end{aligned} \tag{5.16}$$

Equation (5.14) reduces at once to

$$x = c \cos(\lambda t + \phi), \tag{5.17}$$

where c is given by Equation (5.15) and ϕ is determined by Equations (5.16). Since $\lambda = \sqrt{k/m}$, we now write the solution (5.17) in the form

$$x = c \cos\left(\sqrt{\frac{k}{m}}\, t + \phi\right). \tag{5.18}$$

This, then, gives the displacement x of the mass from the equilibrium position O as a function of the time $t (t > 0)$. We see at once that the free, undamped motion of the mass is a *simple harmonic motion*. The constant c is called the *amplitude* of the motion and gives the maximum (positive) displacement of the mass from its equilibrium position. The motion is a *periodic* motion, and the mass oscillates back and forth between $x = c$ and $x = -c$. We have $x = c$ if and only if

$$\sqrt{\frac{k}{m}}\, t + \phi = \pm 2n\pi$$

$n = 0, 1, 2, 3, \ldots; t > 0$. Thus the maximum (positive) displacement occurs if and only if

$$t = \sqrt{\frac{m}{k}}(\pm 2n\pi - \phi) > 0, \tag{5.19}$$

where $n = 0, 1, 2, 3, \ldots$.

The time interval between two successive maxima is called the *period* of the motion. Using (5.19), we see that it is given by

$$\frac{2\pi}{\sqrt{k/m}} = \frac{2\pi}{\lambda}. \tag{5.20}$$

The reciprocal of the period, which gives the number of oscillations per second, is called the *natural frequency* (or simply *frequency*) of the motion. The number ϕ is called the *phase constant* (or *phase angle*). The graph of this motion is shown in Figure 5.2.

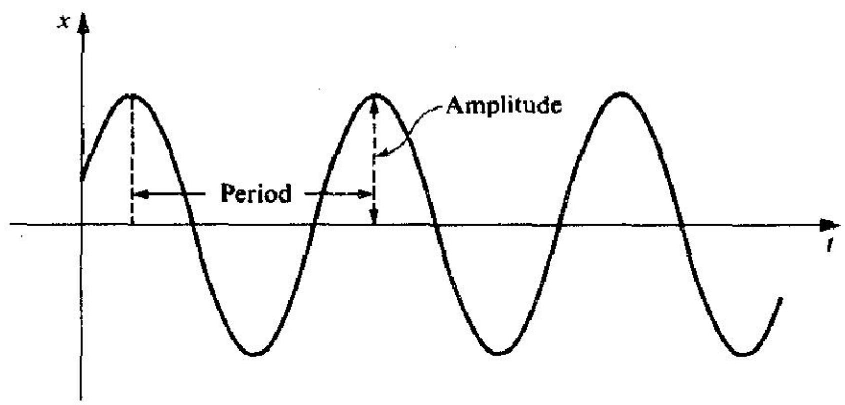


Figure 5.2

5.3 FREE, DAMPED MOTION

We now consider the effect of the resistance of the medium upon the mass on the spring. Still assuming that no external forces are present, this is then the case of *free, damped motion*. Hence with the damping coefficient $a > 0$ and $F(t) = 0$ for all t, the basic differential equation (5.7) reduces to

$$m\frac{d^2x}{dt^2} + a\frac{dx}{dt} + kx = 0. \tag{5.27}$$

Dividing through by m and putting $k/m = \lambda^2$ and $a/m = 2b$ (for convenience) we have the differential equation (5.27) in the form

$$\frac{d^2x}{dt^2} + 2b\frac{dx}{dt} + \lambda^2 x = 0. \tag{5.28}$$

Observe that since a is positive, b is also positive. The auxiliary equation is

$$r^2 + 2br + \lambda^2 = 0. \tag{5.29}$$

Using the quadratic formula we find that the roots of (5.29) are

$$\frac{-2b \pm \sqrt{4b^2 - 4\lambda^2}}{2} = -b \pm \sqrt{b^2 - \lambda^2}. \tag{5.30}$$

Three distinct cases occur, depending upon the nature of these roots, which in turn depends upon the sign of $b^2 - \lambda^2$.

Case 1. Damped, Oscillatory Motion. Here we consider the case in which $b < \lambda$, which implies that $b^2 - \lambda^2 < 0$. Then the roots (5.30) are the conjugate complex numbers $-b \pm \sqrt{\lambda^2 - b^2}\, i$ and the general solution of Equation (5.28) is thus

$$x = e^{-bt}(c_1 \sin\sqrt{\lambda^2 - b^2}\, t + c_2 \cos\sqrt{\lambda^2 - b^2}\, t), \tag{5.31}$$

where c_1 and c_2 are arbitrary constants. We may write this in the alternative form

$$x = ce^{-bt}\cos(\sqrt{\lambda^2 - b^2}\, t + \phi), \tag{5.32}$$

where $c = \sqrt{c_1^2 + c_2^2} > 0$ and ϕ is determined by the equations

$$\frac{c_1}{\sqrt{c_1^2 + c_2^2}} = -\sin\phi,$$
$$\frac{c_2}{\sqrt{c_1^2 + c_2^2}} = \cos\phi.$$

The right member of Equation (5.32) consists of two factors,

$$ce^{-bt} \quad \text{and} \quad \cos(\sqrt{\lambda^2 - b^2}\, t + \phi).$$

The factor ce^{-bt} is called the *damping factor*, or *time-varying amplitude*. Since $c > 0$, it is positive; and since $b > 0$, it tends to zero monotonically as $t \to \infty$. In other words, as time goes on this positive factor becomes smaller and smaller and eventually becomes negligible. The remaining factor, $\cos(\sqrt{\lambda^2 - b^2}\, t + \phi)$, is, of course, of a periodic, oscillatory character; indeed it represents a simple harmonic motion. The product of these two factors, which is precisely the right member of Equation (5.32), therefore represents an oscillatory motion in which the oscillations become successively smaller and smaller. The oscillations are said to be "damped out," and the motion is described as *damped, oscillatory motion*. Of course, the motion is no longer periodic, but the time interval between two successive (positive) maximum displacements is still referred to as the *period*. This is given by

$$\frac{2\pi}{\sqrt{\lambda^2 - b^2}}.$$

The graph of such a motion is shown in Figure 5.5, in which the damping factor ce^{-bt} and its negative are indicated by dashed curves.

The ratio of the amplitude at any time T to that at time

$$T - \frac{2\pi}{\sqrt{\lambda^2 - b^2}}$$

one period before T is the constant

$$\exp\left(-\frac{2\pi b}{\sqrt{\lambda^2 - b^2}}\right).$$

Thus the quantity $2\pi b/\sqrt{\lambda^2 - b^2}$ is the decrease in the logarithm of the amplitude ce^{-bt} over a time interval of one period. It is called the *logarithmic decrement*.

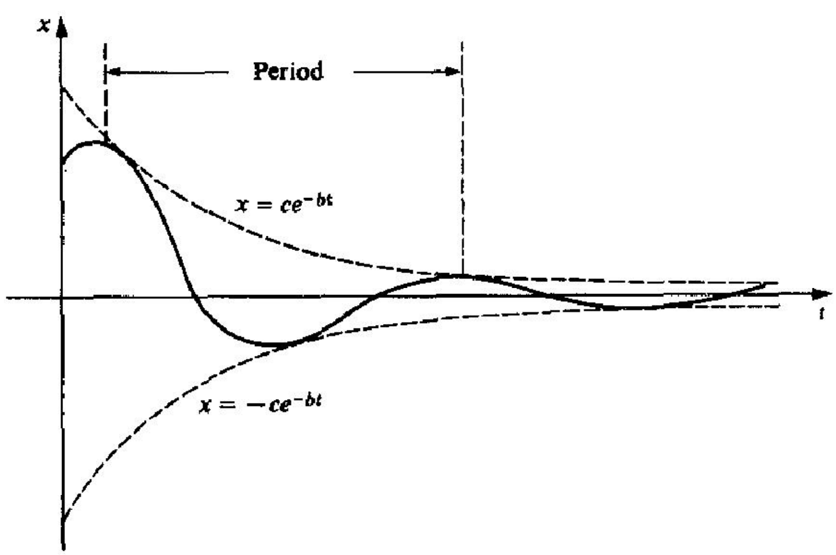


Figure 5.5

If we now return to the original notation of the differential equation (5.27), we see from Equation (5.32) that in terms of the original constants m, a, and k, the general solution of (5.27) is

$$x = ce^{-(a/2m)t}\cos\left(\sqrt{\frac{k}{m} - \frac{a^2}{4m^2}}\, t + \phi\right). \tag{5.33}$$

Since $b < \lambda$ is equivalent to $a/2m < \sqrt{k/m}$, we can say that the general solution of (5.27) is given by (5.33) and that damped, oscillatory motion occurs when $a < 2\sqrt{km}$. The frequency of the oscillations

$$\cos\left(\sqrt{\frac{k}{m} - \frac{a^2}{4m^2}}\, t + \phi\right) \tag{5.34}$$

is

$$\frac{1}{2\pi}\sqrt{\frac{k}{m} - \frac{a^2}{4m^2}}.$$

If damping were not present, a would equal zero and the natural frequency of an undamped system would be $(1/2\pi)\sqrt{k/m}$. Thus the frequency of the oscillations (5.34) in the damped oscillatory motion (5.33) is less than the natural frequency of the corresponding undamped system.

Case 2. Critical Damping. This is the case in which $b = \lambda$, which implies that $b^2 - \lambda^2 = 0$. The roots (5.30) are thus both equal to the real negative number $-b$, and the general solution of Equation (5.28) is thus

$$x = (c_1 + c_2 t)e^{-bt}. \tag{5.35}$$

The motion is no longer oscillatory; the damping is just great enough to prevent oscillations. Any slight decrease in the amount of damping, however, will change the situation back to that of Case 1 and damped oscillatory motion will then occur. Case 2 then is a borderline case; the motion is said to be *critically damped*.

From Equation (5.35) we see that

$$\lim_{t\to\infty} x = \lim_{t\to\infty} \frac{c_1 + c_2 t}{e^{bt}} = 0.$$

Hence the mass tends to its equilibrium position as $t \to \infty$. Depending upon the initial conditions present, the following possibilities can occur in this motion:

1. The mass neither passes through its equilibrium position nor attains an extremum (maximum or minimum) displacement from equilibrium for $t > 0$. It simply approaches its equilibrium position monotonically as $t \to \infty$. (See Figure 5.6*a*.)
2. The mass does not pass through its equilibrium position for $t > 0$, but its displacement from equilibrium attains a single extremum for $t = T_1 > 0$. After this extreme displacement occurs, the mass tends to its equilibrium position monotonically as $t \to \infty$. (See Figure 5.6*b*.)
3. The mass passes through its equilibrium position once at $t = T_2 > 0$ and then attains an extreme displacement at $t = T_3 > T_2$, following which it tends to its equilibrium position monotonically as $t \to \infty$. (See Figure 5.6*c*.)

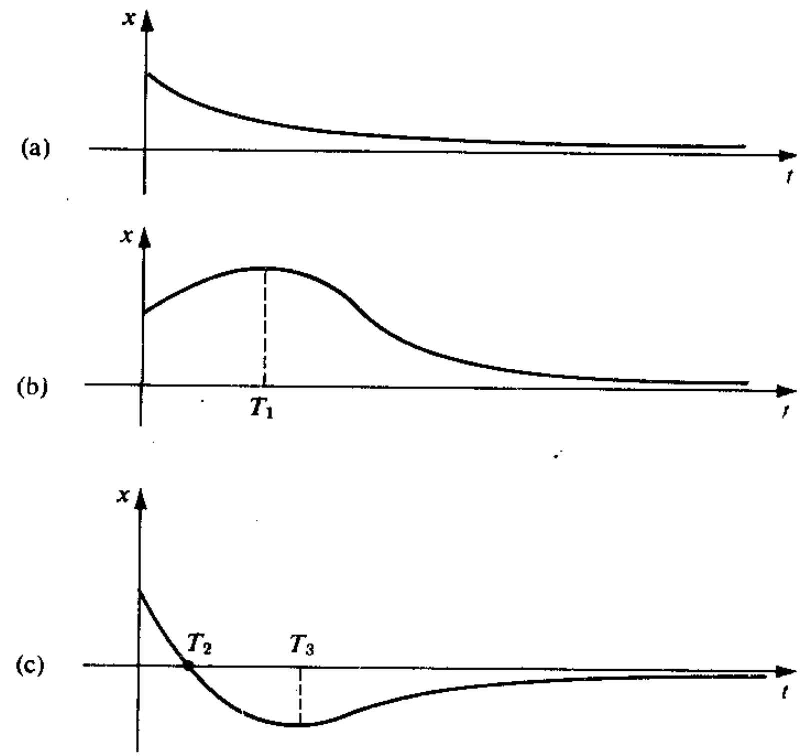


Figure 5.6

Case 3. Overcritical Damping. Finally, we consider here the case in which $b > \lambda$, which implies that $b^2 - \lambda^2 > 0$. Here the roots (5.30) are the distinct, real negative numbers

$$r_1 = -b + \sqrt{b^2 - \lambda^2}$$

and

$$r_2 = -b - \sqrt{b^2 - \lambda^2}.$$

The general solution of (5.28) in this case is

$$x = c_1 e^{r_1 t} + c_2 e^{r_2 t}. \tag{5.36}$$

The damping is now so great that no oscillations can occur. Further, we can no longer say that *every* decrease in the amount of damping will result in oscillations, as we could in Case 2. The motion here is said to be *overcritically damped* (or simply *overdamped*).

Equation (5.36) shows us at once that the displacement x approaches zero as $t \to \infty$. As in Case 2 this approach to zero is monotonic for t sufficiently large. Indeed, the three possible motions in Cases 2 and 3 are qualitatively the same. Thus the three motions illustrated in Figure 5.6 can also serve to illustrate the three types of motion possible in Case 3.

5.4 FORCED MOTION

We now consider an important special case of *forced motion*. That is, we not only consider the effect of damping upon the mass on the spring but also the effect upon it of a periodic external impressed force F defined by $F(t) = F_1 \cos \omega t$ for all $t \geq 0$, where F_1 and ω are constants. Then the basic differential equation (5.7) becomes

$$m \frac{d^2x}{dt^2} + a \frac{dx}{dt} + kx = F_1 \cos \omega t. \tag{5.50}$$

Dividing through by m and letting

$$\frac{a}{m} = 2b, \qquad \frac{k}{m} = \lambda^2, \quad \text{and} \quad \frac{F_1}{m} = E_1,$$

this takes the more convenient form

$$\frac{d^2x}{dt^2} + 2b \frac{dx}{dt} + \lambda^2 x = E_1 \cos \omega t. \tag{5.51}$$

We shall assume that the positive damping constant a is small enough so that the damping is less than critical. In other words we assume that $b < \lambda$. Hence by Equation (5.32) the complementary function of Equation (5.51) can be written

$$x_c = ce^{-bt} \cos(\sqrt{\lambda^2 - b^2}\,t + \phi). \tag{5.52}$$

We shall now find a particular integral of (5.51) by the method of undetermined coefficients. We let

$$x_p = A\cos\omega t + B\sin\omega t. \tag{5.53}$$

Then

$$\frac{dx_p}{dt} = -\omega A\sin\omega t + \omega B\cos\omega t,$$

$$\frac{d^2x_p}{dt^2} = -\omega^2 A\cos\omega t - \omega^2 B\sin\omega t.$$

Substituting into Equation (5.51), we have

$$[-2b\omega A + (\lambda^2-\omega^2)B]\sin\omega t + [(\lambda^2-\omega^2)A + 2b\omega B]\cos\omega t = E_1\cos\omega t.$$

Thus, we have the following two equations from which to determine A and B:

$$-2b\omega A + (\lambda^2-\omega^2)B = 0,$$

$$(\lambda^2-\omega^2)A + 2b\omega B = E_1.$$

Solving these, we obtain

$$A = \frac{E_1(\lambda^2-\omega^2)}{(\lambda^2-\omega^2)^2 + 4b^2\omega^2},$$

$$B = \frac{2b\omega E_1}{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}. \tag{5.54}$$

Substituting these values into Equation (5.53), we obtain a particular integral in the form

$$x_p = \frac{E_1}{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}[(\lambda^2-\omega^2)\cos\omega t + 2b\omega\sin\omega t].$$

We now put this in the alternative "phase angle" form; we write

$$(\lambda^2-\omega^2)\cos\omega t + 2b\omega\sin\omega t$$

$$= \sqrt{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}\left[\frac{\lambda^2-\omega^2}{\sqrt{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}}\cos\omega t + \frac{2b\omega}{\sqrt{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}}\sin\omega t\right]$$

$$= \sqrt{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}\,[\cos\omega t\cos\theta + \sin\omega t\sin\theta]$$

$$= \sqrt{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}\,\cos(\omega t - \theta),$$

where

$$\cos\theta = \frac{\lambda^2-\omega^2}{\sqrt{(\lambda^2-\omega^2)^2 + 4b^2\omega^2}}, \tag{5.55}$$

Thus the particular integral appears in the form

$$x_p = \frac{E_1}{\sqrt{(\lambda^2 - \omega^2)^2 + 4b^2\omega^2}} \cos(\omega t - \theta), \tag{5.56}$$

where θ is determined from Equations (5.55). Thus, using (5.52) and (5.56) the general solution of Equation (5.51) is

$$x = x_c + x_p = ce^{-bt} \cos(\sqrt{\lambda^2 - b^2}t + \phi) + \frac{E_1}{\sqrt{(\lambda^2 - \omega^2)^2 + 4b^2\omega^2}} \cos(\omega t - \theta). \tag{5.57}$$

Observe that this solution is the sum of two terms. The first term, $ce^{-bt} \cos(\sqrt{\lambda^2 - b^2}t + \phi)$, represents the damped oscillation that would be the entire motion of the system if the external force $F_1 \cos \omega t$ were not present. The second term,

$$\frac{E_1}{\sqrt{(\lambda^2 - \omega^2)^2 + 4b^2\omega^2}} \cos(\omega t - \theta),$$

which results from the presence of the external force, represents a simple harmonic motion of period $2\pi/\omega$. Because of the damping factor ce^{-bt} the contribution of the first term will become smaller and smaller as time goes on and will eventually become negligible. The first term is thus called the *transient* term. The second term, however, being a cosine term of constant amplitude, continues to contribute to the motion in a periodic, oscillatory manner. Eventually, the transient term having become relatively small, the entire motion will consist essentially of that given by this second term. This second term is thus called the *steady-state* term.

5.5 RESONANCE PHENOMENA

We now consider the amplitude of the steady-state vibration that results from the periodic external force defined for all t by $F(t) = F_1 \cos \omega t$, where we assume that $F_1 > 0$. For fixed b, λ, and E_1 we see from Equation (5.56) that this is the function f of ω defined by

$$f(\omega) = \frac{E_1}{\sqrt{(\lambda^2 - \omega^2)^2 + 4b^2\omega^2}}. \tag{5.67}$$

If $\omega = 0$, the force $F(t)$ is the constant F_1 and the amplitude $f(\omega)$ has the value $E_1/\lambda^2 > 0$. Also, as $\omega \to \infty$, we see from (5.67) that $f(\omega) \to 0$. Let us consider the function f for $0 < \omega < \infty$. Calculating the derivative $f'(\omega)$ we find that this derivative equals zero only if

$$4\omega[2b^2 - (\lambda^2 - \omega^2)] = 0$$

and hence only if $\omega = 0$ or $\omega = \sqrt{\lambda^2 - 2b^2}$. If $\lambda^2 < 2b^2$, $\sqrt{\lambda^2 - 2b^2}$ is a complex number. Hence in this case f has no extremum for $0 < \omega < \infty$, but rather f decreases monotonically for $0 < \omega < \infty$ from the value E_1/λ^2 at $\omega = 0$ and approaches zero as $\omega \to \infty$. Let us assume that $\lambda^2 > 2b^2$. Then the function f has a relative maximum at $\omega_1 = \sqrt{\lambda^2 - 2b^2}$, and this maximum value is given by

$$f(\omega_1) = \frac{E_1}{\sqrt{(2b^2)^2 + 4b^2(\lambda^2 - 2b^2)}} = \frac{E_1}{2b\sqrt{\lambda^2 - b^2}}.$$

When the frequency of the forcing function $F_1 \cos \omega t$ is such that $\omega = \omega_1$, then the forcing function is said to be in *resonance* with the system. In other words, the forcing function defined by $F_1 \cos \omega t$ is in resonance with the system when ω assumes the value ω_1 at which $f(\omega)$ is a maximum. The value $\omega_1/2\pi$ is called the *resonance frequency* of the system. Note carefully that resonance can occur only if $\lambda^2 > 2b^2$. Since then $\lambda^2 > b^2$, the damping must be less than critical in such a case.

We now return to the original notation of Equation (5.50). In terms of the quantities m, a, k, and F_1 of that equation, the function f is given by

$$f(\omega) = \frac{\dfrac{F_1}{m}}{\sqrt{\left(\dfrac{k}{m} - \omega^2\right)^2 + \left(\dfrac{a}{m}\right)^2 \omega^2}} \tag{5.68}$$

In this original notation the resonance frequency is

$$\frac{1}{2\pi}\sqrt{\frac{k}{m} - \frac{a^2}{2m^2}}. \tag{5.69}$$

Since the frequency of the corresponding free, damped oscillation is

$$\frac{1}{2\pi}\sqrt{\frac{k}{m} - \frac{a^2}{4m^2}},$$

we see that the resonance frequency is less than that of the corresponding free, damped oscillation.

The graph of $f(\omega)$ is called the *resonance curve* of the system. For a given system with fixed m, k, and F_1, there is a resonance curve corresponding to each value of the damping coefficient $a \geq 0$. Let us choose $m = k = F_1 = 1$, for example, and graph the resonance curves corresponding to certain selected values of a. In this case we have

$$f(\omega) = \frac{1}{\sqrt{(1 - \omega^2)^2 + a^2\omega^2}}$$

and the resonance frequency is given by $(1/2\pi)\sqrt{1 - a^2/2}$. The graphs appear in Figure 5.10.

Observe that resonance occurs in this case only if $a < \sqrt{2}$. As a decreases from $\sqrt{2}$ to 0, the value ω_1 at which resonance occurs increases from 0 to 1 and the corresponding maximum value of $f(\omega)$ becomes larger and larger. In the limiting case $a = 0$, the maximum has disappeared and an infinite discontinuity occurs at $\omega = 1$. In this case our solution actually breaks down, for then

$$f(\omega) = \frac{1}{\sqrt{(1 - \omega^2)^2}} = \frac{1}{1 - \omega^2}$$

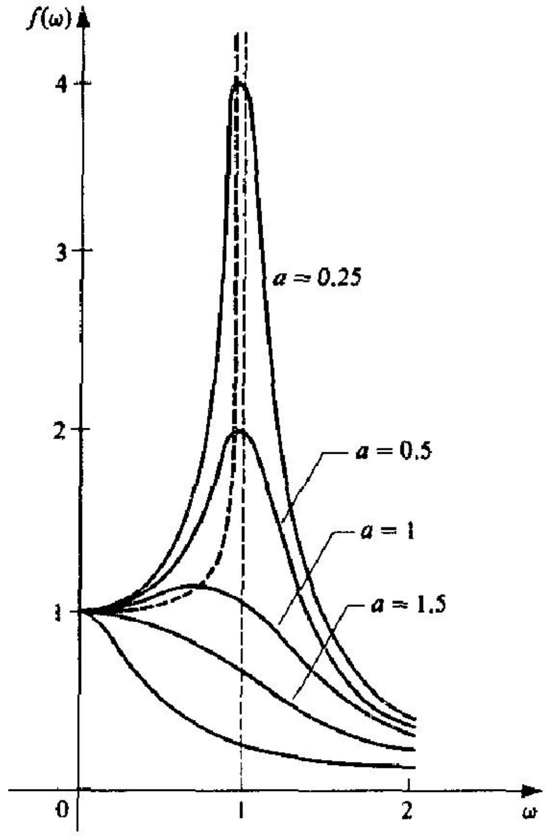


Figure 5.10

and $f(1)$ is undefined. This limiting case is an example of *undamped resonance*, a phenomenon that we shall now investigate.

Undamped resonance occurs when there is no damping and the frequency of the impressed force is equal to the natural frequency of the system. Since in this case $a = 0$ and the frequency $\omega/2\pi$ equals the natural frequency $(1/2\pi)\sqrt{k/m}$, the differential equation (5.50) reduces to

$$m\frac{d^2x}{dt^2} + kx = F_1 \cos\sqrt{\frac{k}{m}}\, t$$

or

$$\frac{d^2x}{dt^2} + \frac{k}{m}x = E_1 \cos\sqrt{\frac{k}{m}}\, t, \tag{5.70}$$

where $E_1 = F_1/m$. Since the complementary function of Equation (5.70) is

$$x_c = c_1 \sin\sqrt{\frac{k}{m}}\, t + c_2 \cos\sqrt{\frac{k}{m}}\, t, \tag{5.71}$$

we cannot assume a particular integral of the form

$$A \sin\sqrt{\frac{k}{m}}\, t + B \cos\sqrt{\frac{k}{m}}\, t.$$

Rather we must assume

$$x_p = At \sin\sqrt{\frac{k}{m}}\,t + Bt\cos\sqrt{\frac{k}{m}}\,t.$$

Differentiating this twice and substituting into Equation (5.70), we find that

$$A = \frac{E_1}{2}\sqrt{\frac{m}{k}} \quad \text{and} \quad B = 0.$$

Thus the particular integral of Equation (5.70) resulting from the forcing function $E_1 \cos\sqrt{k/m}t$ is given by

$$x_p = \frac{E_1}{2}\sqrt{\frac{m}{k}}\,t \sin\sqrt{\frac{k}{m}}\,t.$$

Expressing the complementary function (5.71) in the equivalent "phase-angle" form, we see that the general solution of Equation (5.70) is given by

$$x = c\cos\left(\sqrt{\frac{k}{m}}\,t + \phi\right) + \frac{E_1}{2}\sqrt{\frac{m}{k}}\,t\sin\sqrt{\frac{k}{m}}\,t. \tag{5.72}$$

The motion defined by (5.72) is thus the sum of a periodic term and an oscillatory term whose amplitude $(E_1/2)\sqrt{m/k}t$ increases with t. The graph of the function defined by this latter term,

$$\frac{E_1}{2}\sqrt{\frac{m}{k}}\,t\sin\sqrt{\frac{k}{m}}\,t,$$

appears in Figure 5.11. As t increases, this term clearly dominates the entire motion. One might argue that Equation (5.72) informs us that as $t \to \infty$ the oscillations will become infinite. However, common sense intervenes and convinces us that before this exciting phenomenon can occur the system will break down and then Equation (5.72) will no longer apply!

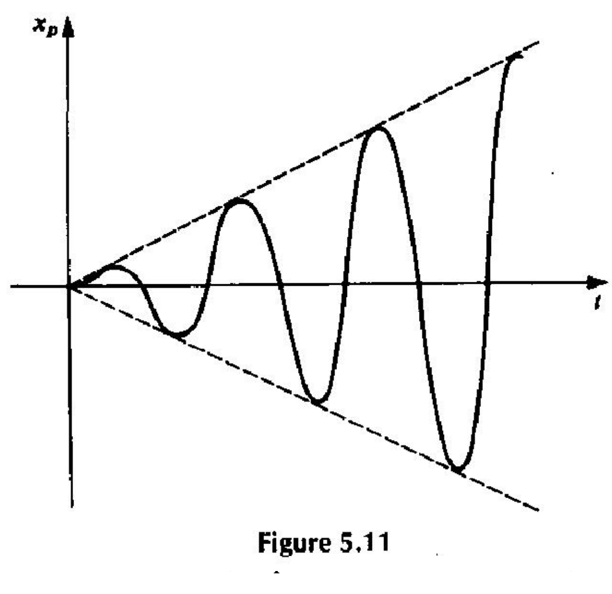


Figure 5.11

5.6 ELECTRIC CIRCUIT PROBLEMS

In this section we consider the application of differential equations to series circuits containing (1) an electromotive force, and (2) resistors, inductors, and capacitors. We assume that the reader is somewhat familiar with these items and so we shall avoid an extensive discussion. Let us simply recall that the electromotive force (for example, a battery or generator) produces a flow of current in a closed circuit and that this current produces a so-called *voltage drop* across each resistor, inductor, and capacitor. Further, the following three laws concerning the voltage drops across these various elements are known to hold:

1. The voltage drop across a resistor is given by

$$E_R = Ri, \tag{5.75}$$

where R is a constant of proportionality called the *resistance*, and i is the current.

2. The voltage drop across an inductor is given by

$$E_L = L\frac{di}{dt}, \tag{5.76}$$

where L is a constant of proportionality called the *inductance*, and i again denotes the current.

3. The voltage drop across a capacitor is given by

$$E_C = \frac{1}{C}q, \tag{5.77}$$

where C is a constant of proportionality called the *capacitance* and q is the instantaneous charge on the capacitor. Since $i = dq/dt$, this is often written as

$$E_C = \frac{1}{C}\int i\,dt.$$

The units in common use are listed in Table 5.1.

TABLE 5.1

Quantity and symbol	*Unit*
emf or voltage E	volt (V)
current i	ampere
charge q	coulomb
resistance R	ohm (Ω)
inductance L	henry (H)
capacitance C	farad

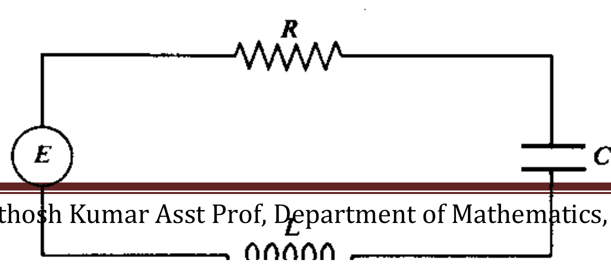
The fundamental law in the study of electric circuits is the following:

Kirchhoff's Voltage Law (Form 1). The algebraic sum of the instantaneous voltage drops around a close circuit in a specific direction is zero.

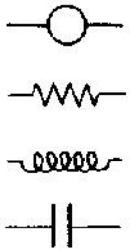
Since voltage drops across resistors, inductors, and capacitors have the opposite sign from voltages arising from electromotive forces, we may state this law in the following alternative form:

Kirchhoff's Voltage Law (Form 2). The sum of the voltage drops across resistors, inductors, and capacitors is equal to the total electromotive force in a closed circuit.

We now consider the circuit shown in Figure 5.12.



Here and in later diagrams the following conventional symbols are employed:



E Electromotive force (battery or generator)

R Resistor

L Inductor

C Capacitor

Let us apply Kirchhoff's law to the circuit of Figure 5.12. Letting E denote the electromotive force, and using the laws 1, 2, and 3 for voltage drops that were given above, we are led at once to the equation

$$L\frac{di}{dt} + Ri + \frac{1}{C}q = E. \tag{5.78}$$

This equation contains *two* dependent variables i and q. However, we recall that these two variables are related to each other by the equation

$$i = \frac{dq}{dt}. \tag{5.79}$$

Using this we may eliminate i from Equation (5.78) and write it in the form

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{1}{C}q = E. \tag{5.80}$$

Equation (5.80) is a second-order linear differential equation in the single dependent variable q. On the other hand, if we differentiate Equation (5.78) with respect to t and make use of (5.79), we may eliminate q from Equation (5.78) and write

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{C}i = \frac{dE}{dt}. \tag{5.81}$$

This is a second-order linear differential equation in the single dependent variable i.

Thus we have the two second-order linear differential equations (5.80) and (5.81) for the charge q and current i, respectively. Further observe that in two very simple cases the problem reduces to a *first*-order linear differential equation. If the circuit contains no capacitor, Equation (5.78) itself reduces directly to

$$L\frac{di}{dt} + Ri = E;$$

while if no inductor is present, Equation (5.80) reduces to

$$R\frac{dq}{dt} + \frac{1}{C}q = E.$$

Before considering examples, we observe an interesting and useful analogy. The differential equation (5.80) for the charge is exactly the same as the differential equation

TABLE 5.2

Mechanical system	*Electrical system*
mass m	inductance L
damping constant a	resistance R
spring constant k	reciprocal of capacitance $1/C$
impressed force $F(t)$	impressed voltage or emf E
displacement x	charge q
velocity $v = dx/dt$	current $i = dq/dt$

Questions	opt1	opt2	opt3	opt4	opt5	opt6	Answer
The magnitude of the force receded to produce a certain elongation of the spring is _____ to the amount of this dlognation, provided this dlognation is not too great	directly proportional	indirectly proportional	eqaul to	greater than			directly proportional
Which of the following is Hooke's law	moduls of (f) = ks	f = ks	f= ma	f not qual to ks			moduls of (f) = ks
The auxiliary equation of d^2 x / dt^2 + λ ^2 x = 0 is _____	r^2 + λ ^2 =0	r^2 + λ ^2 =0	r^2 + λ =0	r + λ =0			r^2 + λ ^2 =0
In Hooke's Law the constant of proportionality can be called as __ constant	Spring	damped	undamped	either damped or undamped			Spring
The Spring at the same time exerts a force upon the mass called the ____ of the spring	acceleration	weight	restoring force	velocity			restoring force
Forces tending to pull the mass downwards are __	negative	positive	zero	Either positive or negative			positive
Forces tending to pull the mass upwards are __	negative	positive	zero	Either positive or negative			negative
The resisting force of the medium is called _______	undamped	Spring	damping force	restoring force			damping force
If F = ma is ______	Newton's first law	Newton's second law	Newton's third law	either second or third law			Newton's second law
If a = 0 then the motion is called _______	undamped	damped	undamped or damped	restoring force			undamped
If a is not eqaul to zero then the motion is called _______	undamped	damped	undamped or damped	restoring force			damped
The differential equation for the motion of the mass on the spring is _	mD^2 / dt ^2 + a dx/dt + kx = F(t)	mD^2 / dt ^2 - a dx/dt + kx = F(t)	mD^2 / dt ^2 - a dx/dt - kx = F(t)	a dx/dt - kx = F(t)			mD^2 / dt ^2 + a dx/dt + kx = F(t)
The number Ǫ is called ______	constant	angle	angle & constant	frequency			angle & constant

The general solution of the differential equation d^2x / dt ^ 2 +64 = 0 is ___	x= C1 sin8t + c2 cos 8t	x= C1 sin8t - c2 cos 8t	x= C1 cos 8t - c2 sin 8t	x= C1 cos 8t + c2 sin 8t			x= C1 sin8t + c2 cos 8t
The auxiliary equation of d^2 x / dt^2 + 2b dx/ dt +λ^2 x = 0 is ______	r^2 - 2br + λ ^2 =0	r^2 + λ =0	r^2 + 2br + λ ^2 =0	r^2 - 2br - λ ^2 =0			r^2 + 2br + λ ^2 =0
The auxiliary equation of d^2 x / dt^2 + 64 x = 0 is _____	r^2 -64 =0	r^2 + 64 =0	- r^2 -64 =0	- r^2 + 64 =0			r^2 + 64 =0
In damped oscillatory motion Ce^- bt is called ___________	damping factor	amplitude	time varying	either varying nor amplitude			damping factor
The natural frequency of an undamped system would be _______	1/2 (k/m)	1/2π (k/m)	1/2π sqrt.(k/m)	1/2π sqrt.(m/k)			1/2π sqrt.(k/m)
In an undamped motion a is __	not eqaul to zero	eqaul to zero	less than zero	greater than zero			eqaul to zero
which of the following is damped oscillator motion is________	b = λ	b > λ	b < λ	b is not equal zero			b < λ
Which of the following is critial damping	b < λ	b > λ	b < λ	b = λ			b > λ
If b^2 - λ^2 < 0 is ____	damped oscillatory motion	critical Damping	over critical damping	damping			damped oscillatory motion
If b^2 - λ^2 = 0 is ____	damped oscillatory motion	critical Damping	over critical damping	damping			critical Damping
If b^2 - λ^2 > 0 is ____	damped oscillatory motion	critical Damping	over critical damping	damping			over critical damping
The general solution of the differential equation d^2x / dt ^ 2 +26dx / dt + λ^2 x = 0 is ___	x= ce^r1t + c2er2t	c=(c1+c2t)e^-bt	c=(c1- c2t)e^-bt	c=(c1-c2t)e^bt			x= ce^r1t + c2er2t
The auxillary equation of d^2x/dt^2 + 4dx/dt +16x =0	r^3+4r^2+16r=0	r^2+4r+16 =0	r+4r^2+16=0	r-4r^2+16=0			r^2+4r+16 =0
The roots of r^2+10r+16 =0 are	r=-2,-8	r=2,8	r=-2,8	r=2,-8			r=-2,-8

The damping is _____ the weight returns to its equilibrium postiom at slower rate	decreased	eqaul	not equal	increased			increased
The roots of r^2+4r+20 = 0 is __	2 + or - 4i	-2 + or - 4i	-2-4	-2-4i			-2 + or - 4i
undamped resonance occurs when there is ___	damping factor	oscillating	no damping	vibration			no damping

Questions	opt1	opt2	opt3	opt4	opt5	opt6	Answer
The string is _______________	flexible	flexible and elastic	elastic	vibration			flexible and elastic
The sring cannot resist __________	vibration	momentum	flexible	bending moment			bending moment
The tension in the ________ is always in the direction of the tangent to the existing profile of the string	string	membrance	vibration	either string nor membrance			string
There is _______ elongation of a single segement of the string	greater	small	no	very small			no
By Hookel's law, the tension is __________	constant	not constant	varible	values			constant
The weight of the string is ___ compared with the tension in the string	equal	small	large	zero			small
The deflection is ___ compared with the length of the string	large	equal	zero	small			small
The slope of the displaced string at any point is ___ compared with unity	equal	small	zero	very small			small
The resultant force is ___ to the mass times the acceleration	equal	large	zero	small			equal
The membrance is ___	flexible	flexible and elastic	elastic	vibration			flexible and elastic
The membrance cannot resist ____	bending moment	momentum	vibration	flexible			bending moment
There is ________ elongation of a single segement of the membrane	large	no	small	verylarge			no
Irrotational motion is	u=1	u=5	u=0	u=-1			u=0

Questions	opt1	opt2	opt3	opt4	opt5	opt6	Answer
The slope of the displaced string at ___ point is small compared with unity	one	two	all	low			two
The Transverse wave velocity is _____	CT	CL	TC	CV			CT
The Laplace equation is also known as ____	Fourier Law	conservation law	Potential law	Burger's law			Potential law
____ can be viewed the special case of heat and wave equation when the dependent variables involved are independent of time	Laplace equation	fouier law	conservation law	Burger's law			Laplace equation
The ___ is a balance between time evolution non-linearity and diffusion	Laplace equation	fouier law	conservation law	Burger's law			Burger's law
Reduced wave equation are also known as ___	Helmholtz equation	Burger's equation	Lamel's equation	potential equation			Helmholtz equation
If the temperature is____heat flows the place of higher temperature to the lower temperature	constant	zero	negative	not constant			not constant
The rate of flow is proportional to the gradient of the temperature is ___ law	conservation	fourier	Potential	burger's			fourier
A force proportional to the product of their masses and inversely proportional to the square of the distance between them is called ___	conservation	fourier	potential	newtons law of gravitation			newtons law of gravitation
Laplace operator may be ____	constant	zero	three dimensional	one only dimensional			three dimensional
If the temperature is not constant ___ flow from place of higher temperature to the lower temperature	energy	heat	cold	potential			heat
The burgers equation is a ___ between time evoluation non linearity and diffusion	balance	unbalance	constant	not constant			balance
CL represents____	potential	transverse wave velocity	gravitational constant	longitudinal wave velocity			longitudinal wave velocity

Questions	opt1	opt2	opt3	opt4	opt5	opt6	Answer
Potential equation is also known as ____	fourier	conservation	laplace	burgers			laplace
laplace equation is also known as ___	fourier	conservation	potential	burgers			potential
If F=GmM/r^2,G is called ____	equlibrium	gravitational	lame's	conservation			gravitational
If F=GmM/r^2,F is called ____	mass	accelaration	force	frequency			force