# CALCULUS

## with analytic geometry

PRENTICE-HALL MATHEMATICS SERIES

DR. ALBERT A. BENNETT, *Editor*

*To G. E. F. Sherwood*

# PREFACE

IN THIS BOOK I have tried to present the ideas and applications of calculus and analytic geometry in a form which is both appealing to students and satisfactory to instructors. The early part of the book is pitched at a level suitable for students with a moderately good knowledge of high school algebra and trigonometry. The interests of students of science and engineering are especially served by the early development of concepts and tools which are needed in the study of physics. No previous study of analytic geometry is required. Instead, topics of analytic geometry are introduced as needed in carrying out the orderly development of the topics in calculus. This arrangement has the twofold advantage of allowing for the introduction of students to some of the fundamental concepts of calculus within the first five or six weeks of the course, and of making it possible to discuss many geometrical topics more effectively with the aid of calculus.

Some people have objected to the current trend toward combining of instruction in analytic geometry and calculus, on the grounds that what is actually happening is that analytic geometry is getting squeezed out. As far as this book is concerned, I believe it can be said accurately of it that the book contains all the material of analytic geometry that students of recent decades have been expected to learn as a prelude to the study of calculus. Knowledge of geometry as raw material for exercises and applications is very important. Also of great importance is the experience of visualizing analytical things in geometric form, and the practice of direct geometric formulation of certain problems. As part of the combination of analytic and geometric method in the applications of calculus, this book employs vector ideas and notation in the discussion of motion, in the treatment of the geometry of lines and planes, and in the discussion of directional derivatives (as components of the gradient).

The treatment in this book of fundamental notions and of theoretical questions is carried out in a manner which I believe to be both understandable by serious students of good average ability, and in reasonable conformity with modern high mathematical standards for precision and rigor. One of the aims of instruction in calculus, apart from its goal of teaching the student techniques for the solution of various important classes of problems, is education of the student in the nature of mathematics as an edifice of logic. The student should learn that definitions are important, and that theorems are logical propositions which are to be demonstrated by reasoning from explicit hypotheses to conclusion. In order to encourage students to reflect on the fundamental concepts of calculus, and on the theoretical development of the subject, I have placed groups of questions on these aspects of the material at the end of every two or three chapters. In my opinion, a test of competence in a substantial course in calculus and analytic geometry should not be based solely on the grounds of formal manipulative skill and ability to apply the technique of the course to stylized problems. This conviction does not mean, however, that I favor a headlong plunge into pure theory. The rich sources of geometry, mechanics, kinematics, and of everyday experience furnish an abundance of problems on which the calculus can demonstrate its power and elegance. Acquaintance with what the calculus can accomplish in a variety of contexts is essential in stimulating student interest and imagination and in providing proper support and background for subsequent educational development.

For the most part, the list of section headings in the table of contents indicates adequately the arrangement of subject matter in the book. Some things deserve special mention.

(1) The law of the mean for derivatives appears very early (at the beginning of Chapter II), and this makes it easier to demonstrate the validity of the claim that the law of the mean is an important instrument of theory.

(2) I have deliberately emphasized Newton's second law and elementary mechanics at several points in the book. Among topics not regularly found in books at this level are the theorem on work and kinetic energy in § 6-9, and the discussion of the principle of motion of the center of mass and the principle of angular momentum in § 20-4. These discussions illuminate the use of calculus in various ways in theoretical mechanics, and I believe it is advantageous to expose students to these applications as they are studying calculus.

(3) The treatment of logarithms is by the integral definition of the natural logarithm function. This is not a new idea, of course, and it has been gaining in favor as compared with the traditional treatment of logarithms as exponents. After long consideration I am convinced of the superiority of the approach via the integral, not only on logical grounds but also for pedagogical reasons. Current high school work in algebra gives a

far too insecure groundwork in exponents to make the traditional definition of logarithms as exponents a comfortable one when it comes to proving anything. Even when allowance is made for postponement of some of the proofs, the discussion of the number $e$ in the traditional approach is cumbersome. The integral treatment of the logarithmic function gives excellent scope for applying the fundamental facts about derivatives and integrals. The whole subject is actually much easier and clearer, and the cumbersome aspects of the discussion of $e$ are entirely avoided. There is, however, need for an adequate motivation of the integral definition of the logarithmic function. This preliminary motivational discussion has been supplied in the present book.

At the root of analysis is the real number system. Some of the basic theory of limits requires only a knowledge of the rules of algebra (the laws governing fields) and the rules governing the use of inequalities. But some of the theory requires knowledge of the "completeness" of the real number system. After early discussion of limits in Chapter I, we return to the subject again in Chapter XIV, partly for a discussion of the theorems about limits of sums, products, and quotients, and partly to lay the groundwork for the study of infinite series. The completeness of the real number system is presented from the Dedekind point of view, in terms of sections. The discussion is held to a very simple level, and only the most essential things are considered. It is not too much to expect students to appreciate such material at this level, if they are given sufficient time and are not tested prematurely.

In conclusion I must express my indebtedness to my now retired colleague Professor Sherwood. Those who are familiar with the book on calculus of which Professor Sherwood and I are co-authors will see much that is familiar in the present book. In some places I have lifted whole blocks of material with little or no change from the Sherwood and Taylor *Calculus*. I have also borrowed freely from our pamphlet *Elementary Differential Equations*. So far as I am aware, the material I have used more or less intact is material which was of my own contribution to the other books. But it is still true that I have been heavily influenced in my views and in my whole experience in teaching calculus by the years of working with Professor Sherwood and using our book in class.

I have also to express my appreciation of the work of Ruthanne Clark, upon whom I depend for superior quality of typing mathematical manuscripts. Lastly, I have been ably assisted by Keith Kendig, a student who has helped greatly in checking the manuscript and preparing the answers.

Here then, is a new book, the fruit of teaching, thinking, and writing spread over a period of nearly five years. It is my attempt to portray how I think calculus and analytic geometry can be taught effectively together.

*Angus E. Taylor*

# CONTENTS

# CHAPTER I

# SLOPES AND RATES OF CHANGE

## 1-1 The Nature of Calculus

In this opening section of the book it is not possible to give a precise and detailed notion of what calculus is as a subject of study. Certain things can be said, however, to indicate what calculus is about and to suggest why it is interesting to study the subject.

The origins of calculus lie in physics and geometry. One branch of physics is concerned with motion, with moving bodies, and with analytical study of the relation between forces applied to bodies and the motion of the bodies under the influence of these forces. This branch of physics is called *mechanics* or *dynamics*. It is of fundamental importance in the applications of physics to engineering. The concept of motion rests essentially on mathematical notions of space and time. To understand motion we must learn what is meant by such terms as *velocity* and *acceleration*, and we must learn how to think and talk precisely about the changes in position of an object (as, for example, a falling stone or a fired bullet) with changes in time. One of the objectives of calculus is to develop the mathematical ideas and tools for understanding and studying motion. Indeed, an exact definition of what is meant by velocity and acceleration is an immediate accompaniment of one of the two main concepts of calculus, that of the *derivative of a function.*

In studying motion it is essential to develop an understanding of certain aspects of geometry. A moving particle traces out a path, which may be a straight line, but which is generally curved. The curve exhibits certain features of the motion, but a full description of the motion requires a correlation between the position of the particle and the time which has elapsed since some initial instant when observation of the motion was begun. It is therefore useful for us to learn how to investigate curves, how to describe them with algebraic formulas or with formulas of types which transcend algebra, and how to discover their properties in detail by examining the formulas. This kind of thing is part of what is called *analytic geometry*.

There is another way in which geometry is related to calculus, quite independently of physics and the concept of motion. In geometry we learn about certain kinds of figures formed by straight lines and planes, and also about certain kinds of curved figures. Triangles, rectangles, polygons, cubes, prisms, and pyramids are examples of figures formed by straight lines and planes. Circles, spheres, cylinders, and cones are examples of curved figures. It is a fundamental matter, in dealing with geometric figures, to know how to calculate circumferences, areas, and volumes. In plane geometry the circle is the simplest curved figure. As we know from high school geometry, the circumference of a circle of radius $r$ is $2\pi r$, and its area is $\pi r^2$; here $\pi$ is a certain number which can be represented *approximately* by the decimal 3.1416. The *precise* decimal representation of $\pi$ does not terminate, and there is no definite pattern of repetition in the digits after the decimal point. These measures of the circumference and area of a circle are arrived at by a method of limits: the circle is regarded as a limit of inscribed or circumscribed polygons. There is a much more general method of limits which may be employed to determine the area of any plane figure bounded by curved lines. The idea of this method is to construct a figure which almost completely fills up the inside of the curved figure, but in such a way that the specially constructed figure is composed entirely of small rectangular pieces, so that its area can be computed simply by adding together the areas of all these pieces. In order to come closer and closer to filling up the curved figure, the sizes of the rectangular pieces must be made smaller and smaller (at least this must be the case for the pieces near the curved edges); the exact area of the curved figure is then obtained by a limiting process (see Fig. 1-1).

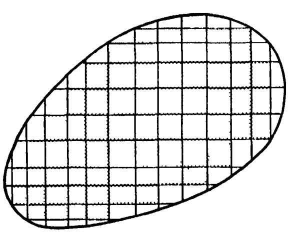


Fig. 1-1

This idea of obtaining the area of a curved figure by a limiting process was used by Archimedes. It is at the root of the concept of the *definite integral of a function*. We have already mentioned that one of the two main concepts of calculus is that of the *derivative* of a function. The other of these two concepts is that of the *integral* of a function. Thus the two principal concepts on which calculus is founded stem, respectively, from the study of motion and the study of areas of curved figures. This is what was meant when, at the outset, we said that the origins of calculus lie in physics and geometry.

The systematic development of calculus began in the 17th century. The English mathematician and physicist Isaac Newton (1642–1727) and the German Gottfried Wilhelm Leibniz (1646–1716) are usually regarded as the founders of systematic method in calculus. But the ideas on which calculus rests had been forming in the minds of other men as well. The ancient Greeks did not use algebraic tools; perhaps but for this fact the Greeks would have achieved what was not achieved until the 17th century. The use of algebra in connection with geometry led to what is now called analytic geometry. In 1637 the French philosopher and mathematician René Descartes (1596–1650) published a famous book which included an important section devoted to the exposition of his ideas about analytic geometry. Eight years earlier another French mathematician, Pierre de Fermat, had developed ideas of much the same kind as those of Descartes. Fermat also investigated methods of determining the position of a line tangent to a curve at a specified point. The basic idea for this is essentially the same as the idea used in determining the velocity of a moving particle.

Since the 17th century the ideas and methods of analytic geometry and calculus have been clarified and improved. But it is interesting that among our very first concerns even today are the applications of calculus to motion problems and to the study of curves.

## 1-2 Fundamentals of Plane Analytic Geometry

In this section we consider *plane* geometry. The ideas of analytic geometry can be applied to solid geometry as well as to plane geometry, however, and later on in the book we shall consider the geometry of three-dimensional space.

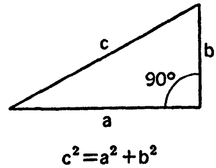



Fig. 1-2

We take for granted that the student has a certain familiarity with plane geometry from his high school studies. Certain propositions and notions of Euclidean geometry are used a great deal in analytic geometry. Of particular importance is the theorem of Pythagoras about the relation between the lengths of the sides of a right triangle (see Fig. 1-2). Also, facts about similar triangles are used a great deal.

### *A Number Scale on a Line*

To get started with the analytic method in geometry we begin by considering the use of numbers to identify points on a single straight line.

Let $L$ be the line and let $O$ and $A$ be two distinct points on $L$. The direction from $O$ toward $A$ is called the *positive direction* on $L$; the opposite direction is called *negative*. Now, using the length $OA$ as the unit of distance, we

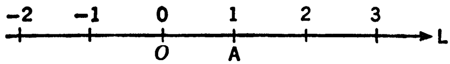



**Fig. 1-3**

set up a one-to-one correspondence between all the points on $L$ and all the real* numbers. The point $O$ corresponds to the number zero, and the point $A$ corresponds to 1. A positive number $p$ corresponds to the point on $L$ that is $p$ units from $O$ in the positive direction, and a negative number $n$ corresponds to the point on $L$ which is $-n$ units from $O$ in the negative direction. It is part of the foundation of Euclidean analytic geometry that this one-to-one correspondence is possible. When this kind of an identification between numbers and points on a line has been established, we say that we have established a *number scale* on the line. The establishment of a number scale commits us to three things: the choice of a positive direction on the line, the location of a zero point (called the *origin* of the scale), and the choice of a unit of length.

*Inequalities and Order*

If the numbers $a$ and $b$ are such that $b - a$ is positive, we indicate this in symbols by writing $a < b$. This is read verbally as "$a$ is less than $b$." Sometimes for emphasis, we say "$a$ is algebraically less than $b$."

**Examples:** $3 < 5$, $-3 < 2$, $-3 < 0$, $-7 < -4$.

Note that "$p$ is positive" and "$0 < p$" mean the same thing, and that "$n < 0$" is the same as "$n$ is negative." The symbolic statement $a < b$ is called an *inequality*. An inequality is equivalent to a statement about relative position of two points on the number scale; $a < b$ means that the direction from $a$ to $b$ on the scale is the positive direction.

Sometimes the inequality symbol is used in the reversed position. Thus $b > a$ means the same as $a < b$. We may read $b > a$ as "$b$ is greater than $a$." If $a < b$ and $b < c$ we sometimes write $a < b < c$. For the points on the number scale this means that $b$ is between $a$ and $c$ and that the positive direction is from $a$ to $b$ to $c$.

The statement "either $a = b$ or $a < b$" is abbreviated symbolically as

* The "real" numbers comprise the positive and negative numbers, and zero. We are in this book mainly concerned with real numbers. But sometimes, as for example, in solving certain quadratic equations, we have occasion to refer to *complex* numbers (sometimes called imaginary numbers). These involve the use of the symbol $i = \sqrt{-1}$. For instance, $i$, $-3i$, $1 - 2i$ and $\frac{1}{2} + \sqrt{2}\, i$ are complex numbers.

$a \leq b$. We may also write it as $b \geq a$. We may read $a \leq b$ as "$a$ is less than or equal to $b$"; and $b \geq a$ may be read as "$b$ is greater than or equal to $a$."

*Absolute Value*

The *absolute value* of a number $c$ is denoted by $|c|$ and defined as follows:

$$|c| = c \text{ if } c \text{ is positive or zero,}$$

$$|c| = -c \text{ if } c \text{ is negative.}$$

Thus the absolute value of a number is never negative, and 0 is the only number whose absolute value is 0.

**Examples:** $|7| = 7$, $|0| = 0$, $|-5| = -(-5) = 5$.

Observe that the absolute value of $c$ is the same as the distance between $c$ and $O$ on the number scale.

If points $P_1$, $P_2$ on $L$ correspond to numbers $x_1$, $x_2$ on the number scale, the distance between $P_1$ and $P_2$ is the absolute value of the difference $x_2 - x_1$; that is, the distance is $x_2 - x_1$ if this difference is positive or zero, and the distance is $x_1 - x_2$ if $x_2 - x_1$ is negative. This is seen to be a correct evaluation of the distance in all cases if we consider separately the three cases: (1) $P_1$ and $P_2$ both in the positive direction from $O$, (2) $P_1$ and $P_2$ both in the negative direction from $O$, (3) $P_1$ and $P_2$ in opposite directions from $O$.

**Examples:** If $x_1 = -3$ and $x_2 = 7$, the distance is $x_2 - x_1 = 10$. If $x_1 = 2$ and $x_2 = -6$, the distance is $x_1 - x_2 = 8$.

*Rectangular Coordinates in a Plane*

Now, turning to the case of a plane, we shall explain how to use pairs of numbers to identify points in the plane. Choose any two perpendicular lines in the plane. Denote by $O$ the intersection of these lines and establish on each line a number scale with origin at $O$. The choice of positive direction on one line can be made independently of the choice on the other line, but we require that the unit of length be the same on the two lines. For convenience of representation on the pages of this book we suppose that these two lines are oriented so that one of them runs across the printed page with the positive direction toward the right, and so that the other line runs up the page with the positive direction toward the top. This orientation is customary, but other orientations are logically permissible and we shall sometimes (later on) use other orientations. Next we assign names to the two lines with their attached number scales. The one extending across the page is called the $x$-axis and the other is called the $y$-axis. If a point $P$ is on the $x$-axis, the number corresponding to it on the number scale is called the *x-coordinate* of $P$. Likewise, a point on the $y$-axis is identified by a certain number, which is called its *y-coordinate.* Now consider any point $P$, anywhere in the plane. Draw a line through $P$ parallel to the $y$-axis (or coinciding with the $y$-axis if $P$ happens to be on that axis). Let $x = a$ be the $x$-coordinate of the point where this line intersects the $x$-axis. Likewise, draw a line through $P$ parallel to (or perhaps coincident with) the $x$-axis, and let $y = b$ be the $y$-coordinate of the point where this line intersects the $y$-axis (see Fig. 1-4). We define the $x$-coordinate and $y$-coordinate of $P$ to be $a$ and $b$, respectively. In referring to the coordinates of $P$ it is convenient to list them as an *ordered pair*, with the $x$-coordinate mentioned first. Thus we say: $P$ has coordinates $(a, b)$. The correspondence between $P$ and its coordinates is a one-to-one correspondence between the totality of points in the plane and the totality of ordered pairs of real numbers. For $P$ determines its coordinates uniquely and each pair of real numbers is the coordinate pair for a uniquely determined point. The point with coordinates $(a, b)$ is often referred to more briefly simply as "the point $(a, b)$."

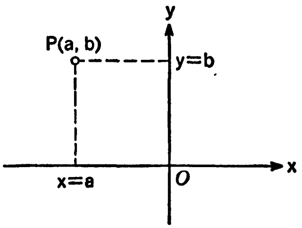



**Fig. 1-4**

When we think of a plane as being provided with an $xy$-coordinate system in the manner here described, we call it *the xy-plane.*

**Example 1:** The point $(4, -2)$ is the point of intersection of the line parallel to the $y$-axis through the point $x = 4$ on the $x$-axis and the line parallel to the $x$-axis through the point $y = -2$ on the $y$-axis (see Fig. 1-5).

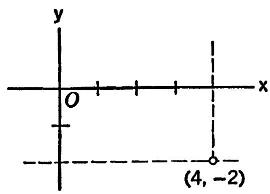



**Fig. 1-5**

The $x$-coordinate of a point is sometimes called its *abscissa,* while the $y$-coordinate is called the *ordinate.* A coordinate system introduced in the manner previously described is called a *rectangular* coordinate system, or a rectangular *Cartesian* coordinate system. The word "Cartesian" is taken from Cartesius, the Latinized name of Descartes.

The basic method of analytic geometry consists in dealing with the coordinates of points instead of with the points. The use of number-pairs to identify points makes it possible to employ arithmetic, algebra, and calculus in the study of geometry.

### *The Distance Formula*

The distance between any two points can be computed in a simple way from the coordinates of these points by an application of the theorem of Pythagoras. Suppose the points are $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$. Denote the distance between the points by $D$. We shall show that $D^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$, so that

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \tag{1}$$

To derive this formula for $D$ we proceed as follows: Construct through $P_1$ a line parallel to the $x$-axis and through $P_2$ a line parallel to the $y$-axis. Let $Q$ be the point where these lines intersect. In general $P_1$, $P_2$, and $Q$ will be distinct points and they will form a right triangle with the right angle at $Q$ (see Fig. 1-6). Therefore $D^2 = \overline{P_1P_2}^2 = \overline{P_1Q}^2 + \overline{QP_2}^2$. But from the construction we see that

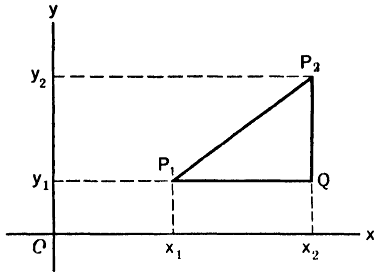



**Fig. 1-6**

$$\overline{P_1Q} = |x_2 - x_1|, \qquad \overline{QP_2} = |y_2 - y_1|.$$

Consequently we have

$$D^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2,$$

and so the formula for $D$ is established. If it happens that $P_1$, $P_2$, and $Q$ are not all distinct, the final formula for $D$ is still valid, though in such a case either $x_2 - x_1 = 0$ or $y_2 - y_1 = 0$, and we do not need the theorem of Pythagoras.

**Example 2:** The distance $D$ between $(4, -1)$ and $(-2, -3)$ is

$$D = \sqrt{(-2 - 4)^2 + (-3 + 1)^2} = \sqrt{40} = 2\sqrt{10}.$$

Observe that, in applying formula (1) to calculate the distance between two points, it does not matter in which order the points are taken.

We take this opportunity to explain an important convention about the use of the square root sign. *In all work in this book, if $A > 0$ then $\sqrt{A}$ denotes the positive square root of $A$.* The negative square root of $A$ is denoted by $-\sqrt{A}$. This convention forces us to write $\sqrt{(-4)^2} = 4$. Thus $\sqrt{a^2} = a$ is correct if $a > 0$, but $\sqrt{a^2} = -a$ if $a < 0$.

*Directed Distances*

When we speak of the distance between two distinct points we shall ordinarily mean a positive number measuring the distance. There are certain occasions, however, in which it is convenient to speak of *directed distances*; which may be negative. If $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ are distinct points on a line parallel to the $x$-axis, we define the directed distance $P_1P_2$ to be $x_2 - x_1$ and the directed distance $P_2P_1$ to be $x_1 - x_2$. Note that we write the points in a definite order, and subtract the coordinate of the first point from that of the second. The directed distance $P_1P_2$ is positive if the direction from $P_1$ to $P_2$ is the same as the positive direction along the $x$-axis. Otherwise the directed distance is negative. Likewise, if $P_1$ and $P_2$ are on a line parallel to the $y$-axis, we define the directed distance $P_1P_2$ to be $y_2 - y_1$ and the directed distance $P_2P_1$ to be $y_1 - y_2$.

**Examples:** Consider $P_1(-1, 2)$, $P_2(3, 2)$, $P_3(3, -3)$, $P_4(-1, -3)$. Then we have directed distances as follows: $P_1P_2 = 4$, $P_2P_3 = -5$, $P_3P_4 = -4$, $P_4P_1 = 5$.

*The Mid-Point Formulas*

It is frequently useful to know the coordinates of the mid-point of the line segment joining two given points. If the points are $(x_1, y_1)$ and $(x_2, y_2)$ the mid-point has coordinates

$$x_0 = \tfrac{1}{2}(x_1 + x_2), \qquad y_0 = \tfrac{1}{2}(y_1 + y_2). \tag{2}$$

To derive these formulas we construct the three lines parallel to the $y$-axis through the respective points $P_1$, $P_2$, and the mid-point $P_0$. These lines cut the $x$-axis at $M_1$, $M_2$, $M_0$, respectively (see Fig. 1-7), and $M_0$ is midway between $M_1$ and $M_2$ (this may be seen by a theorem about transversals cutting the three parallel lines, or more basically, by considerations of similar triangles). But if $M_0$ is midway between $M_1$ and $M_2$ we have

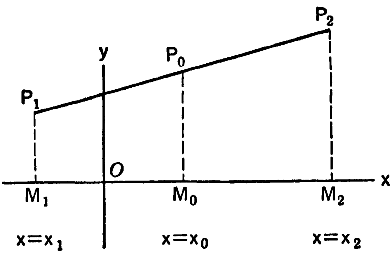


**Fig. 1-7**

$$x_0 - x_1 = x_2 - x_0,$$

and from this it follows that $2x_0 = x_1 + x_2$, or $x_0 = \tfrac{1}{2}(x_1 + x_2)$. The corresponding formula for $y_0$ is obtained by the same kind of argument applied to the $y$-coordinates.

Similar considerations may be employed to find the coordinates of a point $P_0$ on the line joining $P_1$ and $P_2$ such that the distance $P_1P_0$ bears any preassigned ratio to the distance $P_0P_2$. See Exercise 11.

**Example 3:** The four points $P_1$, $P_2$, $P_3$, $P_4$, taken in that order, are

$$(-2, -1), \quad (3, 0), \quad (1, 1), \quad (4, 5).$$

Let $Q_1$ and $Q_2$ be the mid-points of $P_1P_2$ and $P_3P_4$, respectively, and let $R_1$, $R_2$ be the mid-points of $P_2P_3$ and $P_4P_1$, respectively. Show that the mid-point of $Q_1Q_2$ is the same as the mid-point of $R_1R_2$.

Using the mid-point formulas repeatedly, we obtain the coordinates of $Q_1$, $Q_2$, $R_1$, and $R_2$, as follows:

$$Q_1:\ \left(\frac{-2+3}{2}, \frac{-1+0}{2}\right) = \left(\frac{1}{2}, -\frac{1}{2}\right),$$
$$Q_2:\ \left(\frac{1+4}{2}, \frac{1+5}{2}\right) = \left(\frac{5}{2}, 3\right),$$
$$R_1:\ \left(\frac{3+1}{2}, \frac{0+1}{2}\right) = \left(2, \frac{1}{2}\right),$$
$$R_2:\ \left(\frac{4-2}{2}, \frac{5-1}{2}\right) = (1, 2).$$

The mid-point of $Q_1Q_2$ is then $(\frac{3}{2}, \frac{5}{4})$, and the mid-point of $R_1R_2$ is also $(\frac{3}{2}, \frac{5}{4})$.

## EXERCISES

**1.** In the following triangles (determined by the three points listed) find the distance from the first vertex mentioned to the mid-point of the opposite side.
(a) $(-2, -6)$, $(-7, 6)$, $(5, 11)$.
(b) $(-2, 3)$, $(-2, -1)$, $(4, -1)$.
(c) $(2, 6)$, $(-4, 16)$, $(12, 12)$.
(d) $(2, -2)$, $(-1, 3)$, $(-3, 1)$.

**2.** A triangle is determined by each of the following sets of three points. (A) Which triangles are isosceles but not equilateral? (B) Which are equilateral? (C) Which are right triangles?
(a) $(0, 7)$, $(-4, -2)$, $(5, 2)$.
(b) $(4, 5)$, $(-4, -1)$, $(2, -9)$.
(c) $(5, -3)$, $(-7, -5)$, $(-2, 2)$.
(d) $(8, 6)$, $(4, 4)$, $(-1, 10)$.
(e) $(5, 5)$, $(5\sqrt{3}, -5\sqrt{3})$, $(-5, -5)$.
(f) $(1, 6)$, $(-7, -6)$, $(5, -14)$.
(g) $(0, 0)$, $(6, 3)$, $(-2, 4)$.
(h) $(-4, 6)$, $(6, 10)$, $(10, 0)$.

**3.** The points $(0, 0)$, $(5, 2)$, $(8, 7)$, $(3, 5)$ are vertices of a parallelogram. Verify that the diagonals bisect each other by finding the mid-point of each one.

**4.** Plot the points $(4, 1)$, $(1, 3)$, $(-3, 1)$, $(-2, -1)$ and join each point to each of the other points by line segments. Find the lengths of the two segments which intersect.

**5.** Write the proper inequality $<$ or $>$ between the numbers in each of the following pairs, leaving the numbers in the order as given.
(a) $1, -1$; (b) $0, 2$; (c) $-5, -1$; (d) $-5, 2$; (e) $7, 2$; (f) $0, -8$; (g) $-2, 3$.

**6.** Find the directed distances $P_1P_2$ for each of the following ordered pairs of points ($P_1$ is named first). Use the convention described in the text.
(a) $(3, 4)$, $(-1, 4)$; (b) $(3, 4)$, $(3, 9)$; (c) $(-2, 5)$, $(7, 5)$;
(d) $(-1, 6)$, $(-1, -6)$.

**7.** Is it true that $c \leq |c|$ for every real number $c$?

**8.** Do $a \leq b$ and $b < c$ imply that $a < c$?

**9.** Is $-b < -a$ equivalent to $a < b$?

**10.** If $a < b$, demonstrate that $a + c < b + c$ for every choice of $c$. What inequality can you assert about $ac$ and $bc$ if $c \neq 0$?

**11.** (a) Find the coordinates of the point one third the way from (2, 1) to (11, 7) along the line joining these points. Also find the point two thirds of the way from (2, 1) to (11, 7). Use a method similar to that employed in deriving the mid-point formulas.

(b) Derive general formulas for the coordinates of a point $P_0$ if it is on the line joining $P_1$ and $P_2$, and one third the way from $P_1$ to $P_2$. Do likewise for the case in which $P_0$ is two thirds of the way from $P_1$ to $P_2$. The results in the two cases are

$$x_0 = \tfrac{2}{3}x_1 + \tfrac{1}{3}x_2, \qquad x_0 = \tfrac{1}{3}x_1 + \tfrac{2}{3}x_2,$$

with corresponding formulas for $y_0$.

(c) Suppose $P_0$ is on the line joining $P_1$ and $P_2$, and so situated between $P_1$ and $P_2$ that

$$\frac{P_1P_0}{P_0P_2} = \frac{q}{p}.$$

Show that $x_0 = \dfrac{px_1 + qx_2}{p + q}$, $y_0 = \dfrac{py_1 + qy_2}{p + q}$. The formulas in (b) are special cases of this situation.

**12.** (a) For the triangle with vertices $(-3, 5)$, $(5, 2)$, $(9, 8)$ find the point on each median which is two thirds of the way from the vertex to the mid-point of the opposite side. Make the calculations separately for each median and verify that the points found are all the same. Use the results of Exercise 11(b).

(b) Carry out the procedure indicated in part (a) for *any* triangle, denoting the vertices by $P_1(x_1, y_1)$ etc., and show that the coordinates of the point of intersection of the medians are

$$\left[\frac{1}{3}(x_1 + x_2 + x_3), \frac{1}{3}(y_1 + y_2 + y_3)\right].$$

This demonstrates, *by the methods of analytic geometry*, that the medians of a triangle meet in a common point, which is two thirds of the way from a vertex to the mid-point of the opposite side.

**13.** In a modification of the illustrative Example 3 suppose $P_1$ has coordinates $(x_1, y_1)$, and use similar literal coordinates for the other points. Find an expression for the $x$-coordinate of the mid-point of $Q_1Q_2$, and do likewise for the mid-point of $R_1R_2$. Prove in this way that the conclusion reached in Example 3 is valid no matter how the four original points are chosen.

**14.** Does "$|c| < |a|$" mean the same thing as "$-|a| < c$ and $c < |a|$"?

**15.** Demonstrate from the definition of absolute value that $|-c| = |c|$.

**16.** Is it always true that $|ab| = |a|\,|b|$? Justify your answer.

**17.** Under what conditions is it true that $|a + b| = |a| + |b|$? Begin by considering particular numbers, some positive and some negative. Also consider zero. Then try to make a general statement. Do you ever find cases in which $|a + b| > |a| + |b|$? What general conclusion do you draw about $|a + b|$ and $|a| + |b|$?

**18.** Is it always true that $|a| - |b| \leq |a + b|$? Justify your answer.

## 1-3 The Slope of a Line

If a line is not parallel to the $y$-axis there is an important number associated with the line, called its *slope*. This number is defined as follows: Choose any distinct points $P_1$, $P_2$ on the line, with coordinates $(x_1, y_1)$, $(x_2, y_2)$. Draw a line through $P_1$ parallel to the $x$-axis and a line through $P_2$ parallel to the $y$-axis. Let $Q$ be the point of intersection of these lines (see Fig. 1-8). We denote the slope of the line by $m$ and define it as the ratio of two *directed* distances

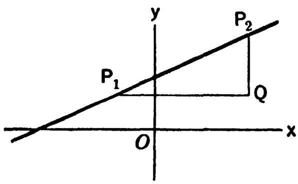



**Fig. 1-8**

$$m = \frac{QP_2}{P_1Q} = \frac{y_2 - y_1}{x_2 - x_1}, \tag{1}$$

i.e., the difference of the two $y$-coordinates divided by the difference of the two $x$-coordinates, both differences being taken with the points in the same order. It is necessary to show that the value of $m$ is the same, no matter how $P_1$ and $P_2$ are chosen on the given line. This is seen to be true by the use of similar triangles. The student should visualize the effect, in Fig. 1-8, of moving either $P_1$ or $P_2$ to a new position on the line.

If we choose $P_2$ so that $x_2 - x_1 = 1$, we see that $m = y_2 - y_1$. Thus the slope is the *algebraic* change in $y$ when $x$ increases by 1 unit and the point $(x, y)$ moves along the line. Observe that the slope is positive if $y$ increases algebraically as $x$ increases, and that the slope is negative if $y$ decreases as $x$ increases. Thus a line rising toward the right has a positive slope and a line rising toward the left has a negative slope. A line parallel to the $x$-axis has slope 0.

### *Angle of Inclination*

The slope of a line can also be thought of as the tangent of a certain angle, called the *angle of inclination* of the line. Let $L$ be any line in the $xy$-plane. If $L$ is parallel to the $x$-axis we define its angle of inclination to be 0°. If $L$ is not parallel to the $x$-axis it intersects the $x$-axis at a point $P_1$. We then define the angle of inclination of $L$ as the *counterclockwise* angle formed at $P_1$ from the positive direction of the $x$-axis to the line $L$. We denote this angle by $\alpha$ (the Greek letter alpha). According to this definition we always have either $\alpha = 0$ or $0 < \alpha < 180$ (assuming $\alpha$ is

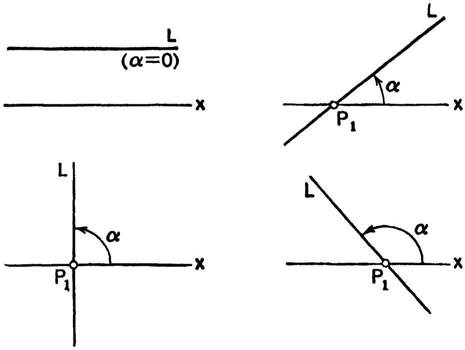



**Fig. 1-9**

reckoned in degrees). See Fig. 1-9 for the inclination in four different positions of $L$. The relation between the slope of a line and its angle of inclination is given by the formula

$$m = \tan \alpha, \tag{2}$$

provided the line is not parallel to the $y$-axis. The truth of this formula is evident from Fig. 1-8 and the definition of $m$ in formula (1); we have only to recall the general definition of the tangent of an angle as a ratio of directed distances.

We have not defined the slope of a line parallel to the $y$-axis. Our definition does not apply to this case, for when a line is parallel to the $y$-axis all points on it have the same $x$-coordinate. If we attempt to use formula (1) for such a line, we find that we cannot, because the denominator of the fraction is 0 (since $x_2 - x_1 = 0$), and division by 0 is impossible. Of course the angle of inclination of such a line is defined; it is 90°, but tan 90° is undefined.

The slope of a line can be computed when we know the coordinates of two points on the line.

**Example 1:** Consider the three points $A(1, -2)$, $B(-2, 3)$, $C(4, 5)$. The slope of the line through $A$ and $B$ is $\frac{3+2}{-2-1} = -\frac{5}{3}$; that of the line through $B$ and $C$ is $\frac{5-3}{4+2} = \frac{1}{3}$, and that of the line through $C$ and $A$ is $\frac{-2-5}{1-4} = \frac{7}{3}$. If we want the angles of inclination we express the slopes in decimal form and use a table of tangents. For instance, the angle of inclination $\alpha$ of $AB$ is determined by the fact that

$$\tan \alpha = -\tfrac{5}{3} = -1.6666\cdots.$$

The negative sign, together with the fact that $\alpha$ must be less than 180°, indicates that $\alpha$ is a second quadrant angle. Since $\tan(180° - \alpha) = -\tan\alpha$ we have

$$\tan(180° - \alpha) = 1.6666\cdots$$
$$180° - \alpha = 59°2'$$
$$\alpha = 120°58' \text{ (approximate).}$$

If a line goes through a specified point with a specified slope, we can easily compute coordinates for a second point on the line, and in this way construct the line.

**Example 2:** A line of slope $\frac{5}{3}$ passes through the point $(-1, 3)$. Find two other points on the line, in opposite directions from the given point.

The slope $\frac{5}{3}$ indicates that if $x$ increases by 3, then $y$ increases by 5. Hence another point on the line is $(-1 + 3, 3 + 5) = (2, 8)$. It is also indicated that if $x$ decreases by 3, then $y$ also decreases by 5. Thus another point on the line is $(-1 - 3, 3 - 5) = (-4, -2)$. These points and the line are shown in Fig. 1-10.

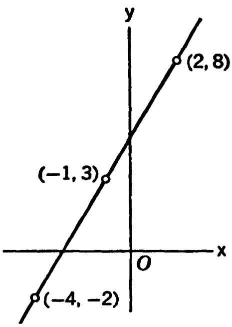



**Fig. 1-10**

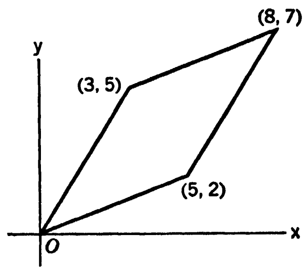



**Fig. 1-11**

*Parallel and Perpendicular Lines*

Two distinct lines are evidently parallel if and only if their angles of inclination are the same. Hence, if they are not parallel to the $y$-axis, *they are parallel if and only if they have the same slope.*

**Example 3:** The four points $(0, 0)$, $(5, 2)$, $(8, 7)$, $(3, 5)$ are consecutive vertices of a parallelogram. For, the slope of the line joining $(0, 0)$ and $(5, 2)$ is $\frac{2}{5}$, that of the line joining $(8, 7)$ and $(3, 5)$ is $\frac{2}{5}$, that of the line joining $(5, 2)$ and $(8, 7)$ is $\frac{5}{3}$, and that of the line joining $(3, 5)$ and $(0, 0)$ is $\frac{5}{3}$. The parallelogram is shown in Fig. 1-11.

Two lines are perpendicular if and only if the angle of inclination of one exceeds that of the other by 90°. Let $\alpha$ and $\beta$ be the angles of inclination, $\alpha$ being the smaller one. The condition for perpendicularity is that $\beta = \alpha + 90°$. One possibility is that $\alpha = 0°$, $\beta = 90°$. This is the special case in which one line is parallel to the $y$-axis. If neither line is parallel to the $y$-axis, the slope of each line is defined, and the condition $\beta = \alpha + 90°$ is equivalent to

$$\tan \beta = \tan (\alpha + 90°).$$

But we know from trigonometry that

$$\tan (\alpha + 90°) = -\text{ctn}\,\alpha = -\frac{1}{\tan \alpha}.$$

*Hence the lines are perpendicular if and only if* $\tan \beta = -\frac{1}{\tan \alpha}$, *i.e., if and only if the slope of one line is the negative reciprocal of the slope of the other.* If we denote the slopes by $m_1$ and $m_2$ (the order of numbering 1 and 2 does not matter), the condition for perpendicularity may be written in either of the forms

$$m_1 = -\frac{1}{m_2}, \qquad m_1 m_2 = -1.$$

These equations apply to the case in which neither slope is 0. If one slope *is* 0, the lines are perpendicular if and only if the other slope is undefined (i.e., the angle of inclination is 90°). The word "normal" is also used with the meaning of "perpendicular." A line perpendicular to another line is said to be normal to it. The normal from a point to a line is the line through the point perpendicular to the line.

**Example 4:** Use slopes to prove that the points $A(6, -5)$, $B(1, 5)$, $C(-2, -1)$ form a right triangle.

The slopes of $AB$, $BC$, and $CA$ are, respectively,

$$\frac{5+5}{1-6} = -2, \quad \frac{-1-5}{-2-1} = 2, \quad \text{and} \quad \frac{-5+1}{6+2} = -\frac{1}{2}.$$

Since 2 and $-\frac{1}{2}$ are negative reciprocals we conclude that the three points form a right triangle with the right angle at $C$ (Fig. 1-12).

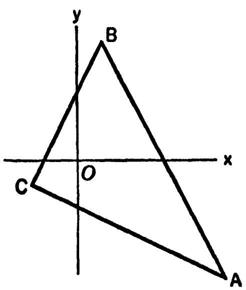



**Fig. 1-12**

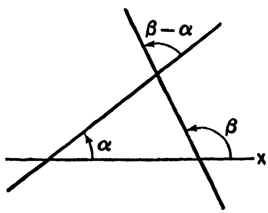



**Fig. 1-13**

*The Angle Between Two Lines*

Slopes can be used to find the angle between two lines. If the angles of inclination are known, the use of slopes is not necessary. In that case we merely subtract the *smaller* angle $\alpha$ from the *larger* angle $\beta$ (see Fig. 1-13).

If the slopes are given and the two lines are not perpendicular we can use the formula

$$\tan(\beta - \alpha) = \frac{\tan\beta - \tan\alpha}{1 + \tan\beta\tan\alpha}$$

to compute the tangent of $\beta - \alpha$. Then $\beta - \alpha$ can be found from a table of tangents. In doing this it should be kept in mind that $\beta - \alpha$ is not as great as 180°.

**Example 5:** If two lines have slopes 2 and $-3$, respectively, find the angle between them.

Clearly the line of slope $-3$ has the greater inclination, so we set $\tan\beta = -3$, $\tan\alpha = 2$. Then

$$\tan(\beta - \alpha) = \frac{-3 - 2}{1 - 6} = 1.$$

We conclude that $\beta - \alpha = 45°$.

*Trigonometry Review*

The general definitions of the trigonometric functions can be made in connection with a rectangular coordinate system. The values of the trigonometric functions are defined as ratios of directed distances. Consider a circle of arbitrary radius $r$ in the $xy$-plane, with center at the origin. If $P$ is a point on this circle it determines an angle $\theta$, *the angle from the positive x-axis to the ray OP* (see Fig. 1-14). For the present we shall employ the *degree* as the unit of angular measurement. There are various ways of assigning the proper number of degrees to the angle $\theta$, depending on conventional agreements which we may make about the generation of the angle. If we regard $\theta$ as generated by the *counterclockwise* rotation of a radius, $\theta$ is positive; if $\theta$ is generated by *clockwise* rotation of a radius, the angle is negative. We may suppose that the rotating radius makes more

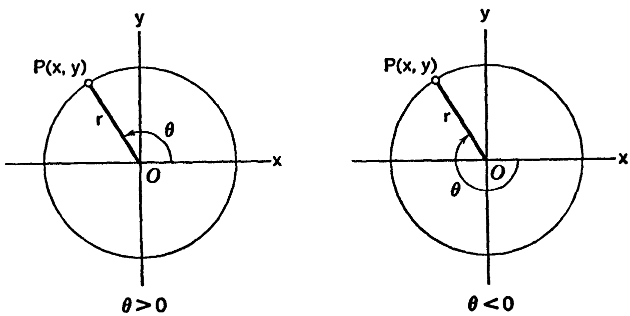



**Fig. 1-14**

than one complete revolution before stopping in the position $OP$; thus $\theta$ may be more than 360° or algebraically less than $-360°$. The value $\theta = 0°$ corresponds to the case when $P$ is on the positive $x$-axis.

The definitions of $\sin\theta$, $\cos\theta$, etc. do not depend on how we imagine the angle $\theta$ to have been generated or upon the units used for measuring $\theta$; *they are completely determined by the position of the point P*. Thus, for instance,

$$\sin 27° = \sin(-333°) = \sin 387°.$$

In terms of the coordinates $(x, y)$ of $P$ and the length $r$ of the radius $OP$ the basic definitions are

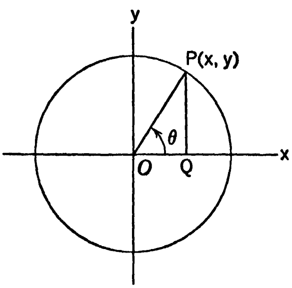



**Fig. 1-15**

$$\sin\theta = \frac{y}{r}, \qquad \csc\theta = \frac{r}{y} = \frac{1}{\sin\theta},$$

$$\cos\theta = \frac{x}{r}, \qquad \sec\theta = \frac{r}{x} = \frac{1}{\cos\theta},$$

$$\tan\theta = \frac{y}{x}, \qquad \text{ctn}\,\theta = \frac{x}{y} = \frac{\cos\theta}{\sin\theta}.$$

Observe that $x$ is the directed distance $OQ$ and that $y$ is the directed distance $QP$ (see Fig. 1-15). Since $r > 0$, $\sin\theta$ and $\cos\theta$ are defined for every possible position of $P$. But $\tan\theta$ and $\sec\theta$ are not defined if $x = 0$, that is, if $P$ is on the $y$-axis, while $\text{ctn}\,\theta$ and $\csc\theta$ are not defined if $y = 0$ ($P$ on the $x$-axis).

The following items are listed for review here because they are relevant to things needed in our study of slopes. The adjacent figures are helpful in expressing the functions of $180° - \theta$ and $90° + \theta$ in terms of functions of $\theta$. On the whole it is better for the student to learn to work out the relations from a figure or from the addition formulas instead of merely memorizing them.

$$\sin(180° - \theta) = \sin\theta,$$
$$\cos(180° - \theta) = -\cos\theta,$$
$$\tan(180° - \theta) = -\tan\theta,$$

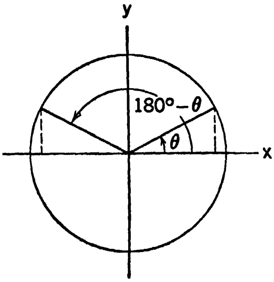



**Fig. 1-16**

$$\sin(90° + \theta) = \cos\theta,$$
$$\cos(90° + \theta) = -\sin\theta,$$
$$\tan(90° + \theta) = -\text{ctn}\,\theta.$$

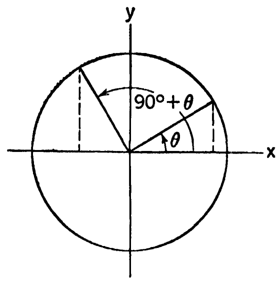



**Fig. 1-17**

The *addition* formulas for the sine and cosine are of fundamental importance:

$$\sin(\theta \pm \phi) = \sin\theta\cos\phi \pm \cos\theta\sin\phi,$$

$$\cos(\theta \pm \phi) = \cos\theta\cos\phi \mp \sin\theta\sin\phi.$$

From these we obtain by division:

$$\tan(\theta \pm \phi) = \frac{\tan\theta \pm \tan\phi}{1 \mp \tan\theta\tan\phi}.$$

It is very convenient to know the values of the trigonometric functions for the angles 0°, 30°, 45°, 60°, 90°. For 0° and 90° we have

$$\sin 0° = \cos 90° = 0, \qquad \sin 90° = \cos 0° = 1.$$

For 30° and 60° the values are easily read off from a 30° − 60° right triangle.

$$\sin 30° = \cos 60° = \frac{1}{2},$$

$$\sin 60° = \cos 30° = \frac{\sqrt{3}}{2},$$

$$\tan 30° = \frac{1}{\sqrt{3}}, \quad \tan 60° = \sqrt{3}.$$

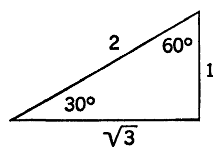



**Fig. 1-18**

For 45° we use an isosceles right triangle.

$$\sin 45° = \cos 45° = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2},$$

$$\tan 45° = 1.$$

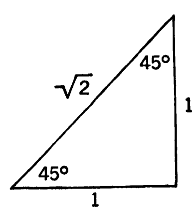



**Fig. 1-19**

## EXERCISES

**1.** Use slopes to decide which of the following sets of three points lie on a straight line.

(a) $(-7, -2), (5, 3), (-\frac{11}{5}, 0)$.
(b) $(-1, -3), (2, -1), (14, 6)$.
(c) $(-2, -9), (3, -1), (8, 7)$.
(d) $(5, 3), (-10, -6), (0, 0)$.
(e) $(-6, 3), (-4, -3), (-2, -10)$.
(f) $(1, 5), (8, 7), (-20, 0)$.

**2.** In each case there is given the slope of a line and a point on it. Find the coordinates of two other points on the line, in opposite directions from the given point.

(a) $\frac{1}{3}, (5, -2)$.
(b) $\frac{2}{3}, (0, -4)$.
(c) $-\frac{2}{5}, (-1, 3)$.
(d) $-4, (-4, 3)$.
(e) $-\frac{1}{2}, (-5, -1)$.
(f) $\frac{3}{2}, (-2, 8)$.

**3.** In each of the following lists of four points call the points $A, B, C, D$ in the order given. (1) Find the cases in which $ABCD$ is a rectangle. (2) Find the cases in which $ABCD$ is a parallelogram but not a rectangle.

(a) $(-6, 3), (8, 7), (12, -5), (-2, -9)$.
(b) $(-15, -13), (-2, -9), (8, 7), (-6, 3)$.
(c) $(1, 5), (3, -1), (-2, -9), (-4, -3)$.
(d) $(-1, -1), (9, 3), (7, 8), (-3, 4)$.
(e) $(3, 6), (-3, 2), (1, -2), (10, 4)$.
(f) $(1, 15), (-5, 13), (1, -5), (7, -3)$.

**4.** (a) Look up in a table the slope of a line whose angle of inclination is: 88°; 89°; 89°30′; 89°59′. (b) What are the slopes corresponding to inclinations of: 92°; 91°; 90°30′?

**5.** Find the interior angles of the triangle with vertices $(4, 5), (-4, 0), (7, -2)$. Draw the figure first. Use tables.

**6.** Proceed as directed in Exercise 5 for the triangle with vertices $(-10, -4), (-4, 6), (5, 3)$.

**7.** What relation is there between the slopes $m_1, m_2$ of lines $L_1, L_2$ if the angles of inclination of these lines are supplementary?

**8.** The line $L$ through $(-3, 0)$ has slope $\frac{1}{2}$ and cuts the $y$-axis at $A$. Find the point $B$ such that $BA$ is perpendicular to $L$ and the line through $B$ parallel to the $y$-axis cuts the $x$-axis at $(-3, 0)$.

**9.** Three vertices $A, B, C$ of a parallelogram $ABCD$ are $(-3, -6), (2, 2), (4, 9)$. Find $D$. Also find a point $E$ so that $AEBC$ is a parallelogram.

**10.** If $(3, -1), (-4, -3), (1, 5)$ are three vertices of a parallelogram, find three different points each of which with the first three will form the vertices of a parallelogram.

**11.** A line has slope $\frac{3}{4}$ and goes through the point $(2, 1)$. (a) Find two points on the line, each 5 units from $(2, 1)$. (b) Find two points on the line, each 8 units from $(2, 1)$.

12. A line has slope $-\frac{5}{12}$ and goes through the point $(-2, 3)$. (a) Find two points on the line, each 5 units from $(-2, 3)$. (b) Find two points on the line, each 13 units from $(-2, 3)$.

13. Find the third vertex of an isosceles triangle whose base is the line joining $(-2, -1)$ and $(6, 3)$, if the altitude on this base is $\sqrt{5}$ (two answers).

14. A right triangle has two of its vertices at $(-3, -4)$ and $(9, 1)$, with the right angle at $(9, 1)$. If the hypotenuse is $\frac{13}{2}\sqrt{5}$ units long, find the other vertex (two answers).

15. A line $L_1$ has slope 2. Lines $L_2$ and $L_3$ have angles of inclination 45° more, and 45° less, respectively, than that of $L_1$. Find the slopes of $L_2$ and $L_3$.

16. (a) Find the slope of a line $L$ if its angle of inclination is 60° less than that of a line of slope $-3$. (b) Find the slope of a line $L$ if its angle of inclination is 30° more than that of a line of slope $-2$.

17. A parallelogram has adjacent edges formed by the lines joining $(0, 0)$, $(a, 0)$ and $(0, 0)$, $(b, c)$, where $a, b, c$ are all positive. (a) Find the remaining vertex of the parallelogram. (b) Under what condition on $a, b, c$ are the diagonals of the parallelogram perpendicular? What is the meaning of this condition as far as the sides of the parallelogram are concerned?

18. Let $P_1$, $P_2$, $P_3$, $P_4$ be any four points. Let $Q_1$ be the mid-point of the segment $P_1P_2$, $Q_2$ the mid-point of $P_2P_3$, $Q_3$ the mid-point of $P_3P_4$, and $Q_4$ the mid-point of $P_4P_1$. Show that $Q_1Q_2$ is parallel to $Q_3Q_4$ and that $Q_2Q_3$ is parallel to $Q_4Q_1$ (i.e., in particular, if we join successively the mid-points of consecutive sides of a quadrilateral, the figure thus formed is a parallelogram).

## 1-4 Equations of Straight Lines

In this section we shall learn how to describe in an algebraic way all the points which lie on a given straight line. This description is made by writing down what is called an *equation of the line.*

First we consider lines parallel to one of the axes. If a line is parallel to the $y$-axis, every point on it has the same $x$-coordinate; if this coordinate is $a$, the equation

$$x = a$$

describes the line in the following sense: *a point* $(x, y)$ *is on the line if and only if* $x = a$. There is no condition placed on $y$ in this case by the demand that $(x, y)$ lie on the line.

Likewise a line parallel to the $x$-axis is characterized by an equation

$$y = b,$$

where $b$ is the $y$-coordinate of every point on the line.

**Example 1:** An equation of the line through $(-3, 0)$ parallel to the $y$-axis is $x = -3$. The equation $y = 1$ describes the line through $(0, 1)$ parallel to the $x$-axis.

If a line is not parallel to either axis it has an equation in which $x$ and $y$ both appear. The equation exhibits the manner in which $x$ and $y$ must be related when $(x, y)$ is a point on the line. Furthermore, this relation between $x$ and $y$ does not hold if $(x, y)$ is not on the line. That is, the validity of the equation for any particular pair $(x, y)$ is both a necessary and sufficient condition for $(x, y)$ to be on the line. Examples will be considered presently. We usually speak of *the* equation of a line, rather than of *an* equation of the line, even though the line may be described by more than one equation. It turns out, however, that any one of these equations can be obtained from any other one simply by multiplying through by some constant factor. For example, $x = -3$ and $2x = -6$ represent the same line. Likewise $x - 2y = 5$ and $3x - 6y = 15$ represent the same line.

The most convenient method for writing down the equation of a line varies slightly, depending on the way in which the information defining the line is furnished.

*The Point-Slope Equation of a Line*

Let $(x_0, y_0)$ be a given point and let $m$ be a given number. *Then the equation*

$$y - y_0 = m(x - x_0) \tag{1}$$

*is the equation of the straight line of slope $m$ through the point $(x_0, y_0)$.* To prove this statement we argue as follows: Suppose $(x, y)$ is any point other than $(x_0, y_0)$ on the line of slope $m$ through $(x_0, y_0)$. Then the slope $m$ is given by the ratio

$$m = \frac{y - y_0}{x - x_0}. \tag{2}$$

If we multiply both sides here by $x - x_0$ we get the equation (1). Conversely, suppose $(x, y)$ is any point such that (1) is true. If $x = x_0$ it follows from (1) that $y = y_0$, and $(x, y)$ coincides with $(x_0, y_0)$. If $x \neq x_0$ we see that (2) follows from (1), so that the line through $(x, y)$ and $(x_0, y_0)$ has slope $m$. We have therefore justified the italicized statement made in connection with (1).

**Example 2:** Find the equation of the perpendicular bisector of the line segment joining the points $(-3, -1)$, $(5, 3)$.

The mid-point of the segment is $(1, 1)$. The slope of the segment is $\frac{1}{2}$, so the slope of the perpendicular bisector is $-2$. The equation of the latter line is therefore

$$y - 1 = -2(x - 1).$$

It can also be written in the form

$$2x + y = 3. \tag{3}$$

*Lines and Linear Equations*

An equation of the form

$$Ax + By + C = 0, \tag{4}$$

where $A$, $B$, and $C$ denote fixed numbers, *with the restriction that $A$ and $B$ are not both 0*, is called a *linear* equation in $x$ and $y$, or an equation *of first degree* in $x$ and $y$. The equations

$$\begin{aligned} 2x + y - 3 &= 0, \\ x + 7 \quad &= 0, \\ y - 5 \quad &= 0, \\ x - 2y \quad &= 0, \end{aligned}$$

are examples of linear equations.

A fundamental fact about straight lines and linear equations is enunciated in the following general theorem:

THEOREM 1-A. *Every straight line in the xy-plane has a linear equation which describes the line. Conversely, every linear equation describes some straight line.*

*Proof.* A line parallel to the $y$-axis has an equation $x = a$, or $x - a = 0$. This is of the form (4), with $A = 1$, $B = 0$, $C = -a$. If the line is not parallel to the $y$-axis it has some slope $m$, and if we choose a point $(x_0, y_0)$ on the line we can write the point-slope equation (1) for the line. This can be put in the form

$$mx - y + y_0 - mx_0 = 0,$$

which is linear. It has the form (4) with $A = m$, $B = -1$, $C = y_0 - mx_0$. We have now verified the truth of the first sentence in the theorem. To finish the proof we must show that every linear equation (4) describes some straight line. We make the proof in two cases. Case 1: $B = 0$. Then $A \neq 0$, since $A$ and $B$ cannot both be 0 in (4). We can then write (4) in the form

$$x = -\frac{C}{A}.$$

This represents a line parallel to the $y$-axis, cutting the $x$-axis at the point $(-C/A, 0)$. Case 2: $B \neq 0$. Now we can write (4) in the form

$$By + C = -Ax,$$

or

$$y + \frac{C}{B} = -\frac{A}{B}(x - 0). \tag{5}$$

If we let $x_0 = 0$, $y_0 = -C/B$, $m = -A/B$, the equation (5) becomes

$$y - y_0 = m(x - x_0).$$

This is the point-slope form. Hence in case 2 our equation (4) represents a line of slope $-A/B$ through the point $(0, -C/B)$. This completes the proof of Theorem 1-A.

**Example 3:** The argument used in the second part of the foregoing proof shows that the equation $8x + 6y - 15 = 0$ represents the line of slope $-\frac{8}{6}$ (or $-\frac{4}{3}$) through the point $(0, \frac{15}{6})$, i.e., $(0, \frac{5}{2})$.

*The Intercepts of a Line*

The $x$-coordinate of the point where a line crosses the $x$-axis is called the *$x$-intercept* of that line. The *$y$-intercept* is the $y$-coordinate of the point where the line cuts the $y$-axis. If we are given the equation of a line which is parallel to neither axis we can easily find both the intercepts.

**Example 4:** Find the intercepts of the line whose equation is $2x - 3y = 4$, and use the information to draw the line.

We set $y = 0$ and solve for $x$:

$$2x = 4, \qquad x = 2.$$

The $x$-intercept is 2. Likewise, setting $x = 0$ and solving for $y$, we obtain

$$-3y = 4, \qquad y = -\tfrac{4}{3}.$$

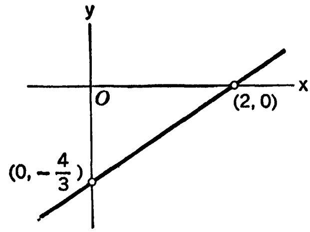


**Fig. 1-20**

The $y$-intercept is $-\frac{4}{3}$. We now know that the line goes through the points $(2, 0)$ and $(0, -\frac{4}{3})$, so we can plot these points and draw the line (see Fig. 1-20).

*The Intercept Form*

Suppose a line has $x$-intercept $a$ and $y$-intercept $b$, so that the line goes through $(a, 0)$ and $(0, b)$. We assume that neither intercept is zero. The slope of the line is

$$m = \frac{b - 0}{0 - a} = -\frac{b}{a},$$

and the equation of the line is

$$y - b = -\frac{b}{a}(x - 0).$$

This can be written as

$$\frac{b}{a}x + y = b,$$

or

$$\frac{x}{a} + \frac{y}{b} = 1. \tag{6}$$

This very symmetrical form of the equation is called the *intercept form.*

*The Slope-Intercept Form*

If a line has slope $m$ and $y$-intercept $b$, its equation is

$$y - b = m(x - 0),$$

or

$$y = mx + b. \tag{7}$$

This is called the *slope-intercept* form.

**Example 5:** Transform the equation $7x + 8y + 5 = 0$ to the slope-intercept form, and use the result to write the equation of the line perpendicular to the first one, with the same $y$-intercept.

To put the equation in the required form we solve for $y$:

$$y = -\tfrac{7}{8}x - \tfrac{5}{8}.$$

Then the coefficient of $x$ is the slope, and the constant term is the $y$-intercept:

$$m = -\tfrac{7}{8}, \qquad b = -\tfrac{5}{8}.$$

The slope of the required perpendicular is $\tfrac{8}{7}$, so its equation is

$$y = \tfrac{8}{7}x - \tfrac{5}{8}.$$

If we wish, this can be written in the form

$$64x - 56y = 35.$$

*The Intersection of Two Lines*

Two distinct lines are either parallel or they intersect in just one point. If this point is $(x_1, y_1)$, the pair $(x_1, y_1)$ satisfies the equation of each line, and it is the only pair which satisfies both equations. *If the equations are given, we find the point of intersection by solving the two equations as a pair of simultaneous linear equations.*

**Example 6:** Find the point of intersection of the lines $2x + 5y = 4$, $3x - 4y + 17 = 0$.

We solve simultaneously by elimination:

$$\begin{array}{rcl} 8x + 20y &=& 16 \\ \underline{15x - 20y} &=& \underline{-85} \\ 23x &=& -69, \qquad x = -3, \end{array}$$

$$4y = 3x + 17 = -9 + 17 = 8, \qquad y = 2.$$

The point of intersection is $(-3, 2)$ (see Fig. 1-21).

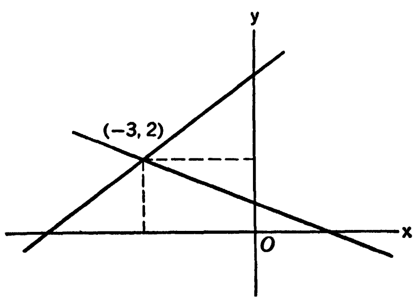



**Fig. 1-21**

The work can be checked graphically by drawing the lines. Gross errors can be detected in this way.

*Parallel Lines*

If two distinct lines are parallel, there is no point of intersection. The parallelism may be detected by examining the equations of the lines. Let the equations be

$$A_1x + B_1y + C_1 = 0,$$

and

$$A_2x + B_2y + C_2 = 0.$$

These lines are *identical* if and only if the two sets of numbers

$$A_1, B_1, C_1 \quad \text{and} \quad A_2, B_2, C_2$$

are proportional. They are distinct and parallel if and only if the pairs

$$A_1, B_1 \quad \text{and} \quad A_2, B_2$$

are proportional, but the proportionality does not extend to $C_1$ and $C_2$.

**Example 7:** The lines

$$3x - 4y + 1 = 0 \quad \text{and} \quad 6x - 8y + 9 = 0$$

are distinct and parallel. The lines

$$2x + 5y = 4 \quad \text{and} \quad 6x + 15y = 12$$

are identical.

## EXERCISES

**1.** Draw a figure for each part of the exercise. Find the equation of the line
(a) through (1, 3) with slope $-2$;
(b) with $x$ and $y$ intercepts $-1$ and 4, respectively;
(c) with $y$-intercept 6 and slope 3;
(d) through (3, 6) parallel to $-2x + 5y = 7$;
(e) through (6, $-2$) perpendicular to $2x + 5y = 3$;
(f) through (3, 4) and (5, $-2$);
(g) with slope $-\frac{1}{2}$ and $x$-intercept 4;
(h) through (2, $-3$) with an inclination of 135°;
(i) through (0, 5), with positive slope, and forming with the axes a triangle of area 20 square units.

**2.** Read off the slope and the $y$-intercept after putting the equation in slope-intercept form. Draw a figure in each case.

(a) $2x + y - 7 = 0$.
(b) $4x + 7y + 3 = 0$.
(c) $2x + 5y = 11$.
(d) $\frac{x}{7} + \frac{y}{3} = 1$.
(e) $\frac{x}{5} - \frac{y}{6} = 1$.
(f) $3x - y + 6 = 0$.
(g) $x = -y + 7$.
(h) $x = 2y$.
(i) $x = 3y + 6$.
(j) $\frac{y}{4} - \frac{x}{3} = \frac{1}{6}$.

**3.** Put each equation in intercept form, and draw the corresponding line.

(a) $x - 2y = 4$.
(b) $3x + 2y = 12$.
(c) $5x - 3y + 15 = 0$.
(d) $3y - 4x = 12$.
(e) $x = 3y - 5$.
(f) $3y = 4x + 7$.

**4.** What does each statement imply about $A$, $B$, $C$ in the line whose equation is $Ax + By + C = 0$?

(a) The slope is $\frac{3}{2}$.
(b) The $x$- and $y$-intercepts are 4 and 3, respectively.
(c) The line goes through the origin.
(d) The line goes through $(1, 1)$.
(e) The line is parallel to the $x$-axis.
(f) The line is perpendicular to the $x$-axis.
(g) The line is parallel to $3y = 2x - 4$.
(h) The line is perpendicular to $2x - 5y = 7$.
(i) The line is identical with $y = 3x - 4$.

**5.** Draw a figure for each part of the exercise. Find the equation of the line:

(a) through $(-2, 1)$ if a line perpendicular to the line has inclination 120°;
(b) through $(-3, -4)$ and tangent to the circle of radius 5 with center at the origin;
(c) perpendicular to the line segment joining $(-1, -3)$ and $(2, -5)$, at its mid-point;
(d) through $(1, 1)$ and the intersection of the lines $x + y - 6 = 0$, $x - 2y + 6 = 0$;
(e) with $y$-intercept 6 and perpendicular to the line through $(1, 1)$ and $(-7, 7)$;
(f) with $x$-intercept 3 and parallel to the line with $x$ and $y$ intercepts 6 and $-4$, respectively;
(g) perpendicular to the line $x - 3y = 6$ at the mid-point of the segment cut from this line by the axes.

**6.** In the following list of pairs of straight lines find the point of intersection in each case in which there is a unique such point. If the lines are coincident, or distinct and parallel, state this fact.

(a) $x + 2y - 3 = 0$, $4x - y - 3 = 0$.
(b) $2x - y + 7 = 0$, $4x = 2y + 3$.
(c) $3x + y - 2 = 0$, $4x + 7y + 3 = 0$.
(d) $x - \frac{1}{2}y = 1$, $2x = y + 2$.
(e) $3x - 5y - 10 = 0$, $x + y + 1 = 0$.
(f) $y = \frac{4}{3}x - 5$, $6y - 8x = 1$.

**7.** (a) If $x$ denotes temperature in degrees Fahrenheit and $y$ denotes temperature in degrees centigrade, find the equation connecting $x$ and $y$, given that it is linear and that we have the following correspondences:

| | *Fahrenheit* | *Centigrade* |
|---|---|---|
| Melting ice........ | 32° | 0° |
| Boiling water...... | 212° | 100° |

(b) Draw the line represented by the equation. What is its slope?
(c) What statement about temperature expresses the value of the $y$-intercept?
(d) What temperature is the same on the Fahrenheit and centigrade scales?
(e) What is normal body temperature (98.6°F) on the centigrade scale?
(f) In some European countries 18°C is called "room temperature." What is this on the Fahrenheit scale?

**8.** The three given equations define the sides of a triangle. Find the common point of intersection of the altitudes produced (lines through a vertex, perpendicular to the opposite side).

$$\begin{aligned} 7x - 12y &= 42, \\ 7x + 20y &= 98, \\ 21x - 10y &= -56. \end{aligned}$$

**9.** Find the common point of intersection of the perpendicular bisectors of the sides of the triangle, the equations of whose sides are $4x - 3y + 30 = 0$, $x + y = 10$, $4x + 25y + 86 = 0$.

**10.** If any triangle is given, coordinate axes may be chosen so that one vertex is on the positive $y$-axis, say at $(0, b)$, while the other vertices are on the $x$-axis, say at $(a, 0)$ and $(c, 0)$. (a) Find the equations of the perpendicular bisectors of the sides, and find their common point of intersection. (b) Find the equations of the altitudes produced, and locate their common point of intersection.

## 1-5 Graphs and Equations

In the case of straight lines we have seen that every linear equation in $x$ and $y$ represents some straight line, and that every straight line has a linear equation. This correspondence between a geometrical configuration (in this case a line) and an equation which describes it exists in the case of many other types of geometrical configuration. For instance, a circle can be described by an equation.

**Example 1:** The circle of radius 2 with center at the origin is described by the equation

$$x^2 + y^2 = 4. \tag{1}$$

For, a point $(x, y)$ is on the specified circle if and only if the distance between $(x, y)$ and $(0, 0)$ is 2, that is, if and only if

$$\sqrt{(x - 0)^2 + (y - 0)^2} = 2,$$

or $x^2 + y^2 = 4$.

**Example 2:** In the same way we see that the circle with radius $r$ and center $(a, b)$ is described by the equation

$$(x - a)^2 + (y - b)^2 = r^2. \tag{2}$$

Various other types of curves may be defined by geometrical conditions which, when expressed in terms of coordinates, yield an equation which describes the curve. The plan of this book does not call for detailed consideration of such matters now, but we shall discuss one example.

**Example 3:** Find the equation of the curve which is composed of all points $(x, y)$ such that the distance from $(x, y)$ to $(0, \frac{1}{4})$ is the same as the perpendicular distance from $(x, y)$ to the line $y = -\frac{1}{4}$.

It is evident from a diagram that the point $(x, y)$ cannot be below the $x$-axis if it is to satisfy the specified condition. Hence in what follows we shall assume that $y \geq 0$. The distance from $(x, y)$ to $(0, \frac{1}{4})$ is

$$\sqrt{(x-0)^2 + (y-\tfrac{1}{4})^2} = \sqrt{x^2 + y^2 - \tfrac{1}{2}y + \tfrac{1}{16}}.$$

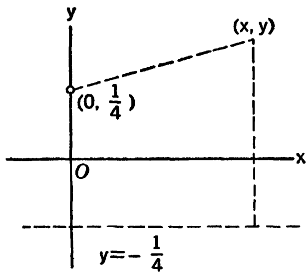



**Fig. 1-22**

The perpendicular distance from $(x, y)$ to the line $y = -\frac{1}{4}$ is $y + \frac{1}{4}$ (see Fig. 1-22). The geometrical condition on $(x, y)$ is therefore expressed by the equation

$$\sqrt{x^2 + y^2 - \tfrac{1}{2}y + \tfrac{1}{16}} = y + \tfrac{1}{4}. \tag{3}$$

This finishes the problem in a certain sense, but the equation (3) can be put into a simpler equivalent form by squaring both sides:

$$x^2 + y^2 - \tfrac{1}{2}y + \tfrac{1}{16} = y^2 + \tfrac{1}{2}y + \tfrac{1}{16}. \tag{4}$$

On transposing and cancelling we obtain the equation

$$y = x^2. \tag{5}$$

This equation is *equivalent* to (3); that is, a point $(x, y)$ satisfies (5) if and only if it satisfies (3). We have already shown that (3) implies (5), so all that is needed is to show that (5) implies (3). Now (5) implies (4), for (4) can be obtained from (5) merely by adding $y^2 - \frac{1}{2}y + \frac{1}{16}$ to each side. Since $y^2 + \frac{1}{2}y + \frac{1}{16} = (y + \frac{1}{4})^2$, (4) implies either (3) or the following equation:

$$\sqrt{x^2 + y^2 - \tfrac{1}{2}y + \tfrac{1}{16}} = -(y + \tfrac{1}{4}). \tag{6}$$

But if (6) were true we should have $-y - \frac{1}{4} \geq 0$, or $y \leq -\frac{1}{4}$, since the radical sign denotes a nonnegative square root. But (5) implies that $y \geq 0$. Hence (6) cannot hold if (5) does; therefore (5) implies (3).

The set (i.e., assemblage or collection) of all points $(x, y)$ which satisfy a given equation in $x$ and $y$ is called the *graph* of the equation. For instance, the graph of (1) is the circle described in Example 1. The graph of a linear equation in $x$ and $y$ is the straight line corresponding to that equation. The graph of (5) is the curve described in Example 3.

Some idea of the appearance of the graph of an equation can be obtained by plotting points which satisfy the equation. If a judicious choice is made of the points which are plotted, a careful examination of the facts revealed by the equation itself may enable us to construct a reasonably accurate freehand drawing representing the graph. We shall in the course of this book develop techniques for detecting the essential features of a graph by examining its equation. Just now we begin with the rather simple task of constructing the graph of the equation which was obtained in Example 3.

**Example 4:** Discussion of the graph of $y = x^2$.

The graph consists of all points $(x, x^2)$, where $x$ can be assigned any value. A partial table of such points can be made as follows:

| $x$ | $0$ | $\pm\frac{1}{4}$ | $\pm\frac{1}{2}$ | $\pm\frac{3}{4}$ | $\pm 1$ | $\pm\frac{3}{2}$ | $\pm 2$ |
|---|---|---|---|---|---|---|---|
| $y$ | $0$ | $\frac{1}{16}$ | $\frac{1}{4}$ | $\frac{9}{16}$ | $1$ | $\frac{9}{4}$ | $4$ |

Several features can be observed at once: The $y$-coordinate is never negative. For every $y$ except 0 there are two values of $x$, one positive and one negative

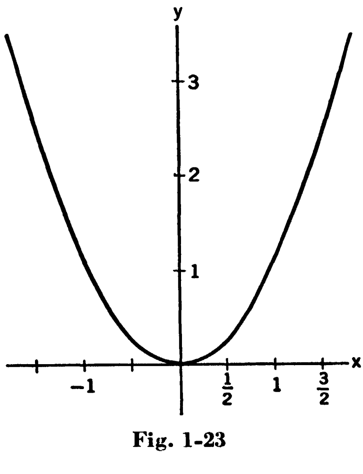



**Fig. 1-23**

but with the same absolute value. This means that points on the graph can be arranged in pairs which are situated symmetrically with respect to the $y$-axis. When $0 < x < 1$ we have $0 < y < x$ and when $1 < x$ we have $x < y$ (squaring a positive number decreases it if the number is less than 1 and increases it if the number is greater than 1). As $x$ gets large $y$ gets large also, and it becomes much larger than $x$. The graph is shown in Fig. 1-23. The curve is called a *parabola*. We shall learn much about parabolas later on.

In constructing the graph of an equation it is important to bear in mind that it is likely to be a matter of greater interest to know in a general way what the graph looks like than to know indiscriminately a large number of points on the graph. It is therefore worth while learning how to discover the most important features of a graph without wasted effort in plotting too many points.

### EXERCISES

**1.** Draw a figure for each part of the exercise. Find the equation of the circle
(a) of radius 1 with center at the point $(1, 0)$;
(b) with center at $(1, 1)$, if the circle goes through the origin;
(c) with center at $(-2, 3)$, tangent to the $y$-axis;
(d) with center at $(3, 2)$, tangent to the line $y = -1$;
(e) tangent to the $x$-axis, the line $x = 10$ and the line $x = 2$ (two possibilities);
(f) through the three points $(1, 0)$, $(7, 0)$, $(4, -1)$.

**2.** (a) Find the equation whose graph consists of all points $(x, y)$ such that the distance from $(x, y)$ to $(1, 0)$ is the same as the distance from $(x, y)$ to the line $x = -1$. (b) Draw the graph.

**3.** (a) Find the equation whose graph consists of all points equidistant from the point $(0, -\frac{1}{2})$ and the line $y = \frac{1}{2}$. (b) Draw the graph.

**4.** Make graphs of $y = x^2$, $y = x^3$ and $y = x^4$ *on the same coordinate axes*, using 10 centimeters as the unit of length, and paying primary attention to the values of $x$ between 0 and 1.2.

**5.** Construct the graphs of each of the following equations. Use symmetry as much as possible. Make a fairly large-scale freehand drawing based on a reasonable number of well distributed points.

| | |
|---|---|
| (a) $y = x^3$. | (e) $y + 2 = \frac{1}{8}x^2$. |
| (b) $y^3 = x$. | (f) $4x^2 + y = 4$. |
| (c) $x^2 = y$. | (g) $y^2 = 4(x + 1)$. |
| (d) $y^2 = x^3$. | (h) $16x + y^2 = 16$. |

## 1-6 Functions

In the applications of mathematics there are many instances of situations in which one thing is said to be a function of another:

(a) The area $A$ of a square is a function of the length $x$ of a side of the square.

(b) The volume $V$ of a sphere is a function of the radius $r$.

(c) The sine $S$ of an angle of $\theta$ degrees is a function of the angle.

(d) The base 10 logarithm $y$ of a number $x$ is a function of the number.

(e) If a baseball is thrown straight up, the height $h$ to which it will rise is a function of the initial velocity $v$.

(f) If a stone is dropped over a vertical cliff, the distance $s$ it falls in $t$ seconds is a function of $t$.

In all examples of this kind there are certain common features, and it is these common features which furnish the basis for the general function concept. The most striking feature is that we have a *pairing* of things, and that the numerical measure of a certain one of the things is determined by the numerical measure of the other. In these particular examples the pairings and the way in which one thing depends upon the other are as follows:

(a) $(x, A)$, $A = x^2$.
(b) $(r, V)$, $V = \frac{4}{3}\pi r^3$.
(c) $(\theta, S)$, $S = \sin\theta$.
(d) $(x, y)$, $y = \log_{10} x$.
(e) $(v, h)$, $h = v^2/64$ ($h$ in feet, $v$ in feet per second).
(f) $(t, s)$, $s = 16t^2$ ($s$ in feet, $t$ in seconds).

The formulas in (e) and (f) come from the laws of freely falling bodies. We do not bother to explain the derivation of these formulas at this time. In the pairings, the first named quantity can be assigned various values; the value of the second quantity is then determined.

Our general definition of a function, for the purposes of elementary calculus and analytic geometry, is as follows: A function is a definitely specified collection of ordered number pairs of such a nature that, if we symbolize the pairing by $(x, y)$, there is a unique value of $y$ corresponding to each allowable value of $x$. In other words, we get all pairs by assigning to $x$ all possible values in a certain preassigned collection of numbers, and then pairing with that $x$ the uniquely determined value of $y$ that goes with it.

**Example 1:** The base 10 logarithm function is the collection of all pairs $(x, \log_{10} x)$, where $x$ varies over all positive numbers.

**Example 2:** The square root function is the collection of all pairs $(x, \sqrt{x})$, where $x$ can be any nonnegative number and $\sqrt{x}$ denotes the nonnegative square root of $x$.

**Example 3:** The formula $y = \sqrt{1 - x^2}$ defines a function, namely, the collection of all pairs $(x, y)$, where $x$ can be assigned any value such that $-1 \le x \le 1$, and for each $x$ the corresponding $y$ is $y = \sqrt{1 - x^2}$.

Of course there is nothing essential about the choice of the letters $x, y$ in the definition of a function. Thus the formula $h = v^2/64$ in the foregoing illustration (e), defines a function consisting of all pairs $(v, v^2/64)$. If $h$ and $v$ are to be interpreted in the manner stated, the allowable values of $v$ are all positive. But if we ignore physical interpretations it is clear that the formula $h = v^2/64$ assigns a unique value to $h$, no matter what value of $v$ is given.

If the pairs constituting a function are symbolized by $(x, y)$, $x$ is called the *independent variable* and $y$ is called the *dependent variable.* The set (collection) of allowable values of $x$ is called the *domain of definition* of the function, and the set of corresponding values of $y$ is called the *range of values* of the function. It has long been customary to say that the dependent variable "is a function of" the independent variable, and we shall sometimes follow this custom. It must be realized, however, that the dependent variable is not actually itself the function. The function is the collection of all pairs $(x, y)$. It is this collection which exhibits the way in which $y$ depends on $x$. In some discussions of the function concept the rule, or law of correspondence, by which $y$ depends on $x$ is called the function. The law of correspondence determines the collection of pairs $(x, y)$, and the collection of pairs exhibits the law of correspondence. Hence there is nothing more fundamental than a matter of nomenclature to distinguish between these usages of the word "function." The collection of pairs $(x, y)$ seems more tangible than the law of correspondence, and it is perhaps on that account that the definition of a function is founded on the "collection of pairs" concept.

The rule which generates a function is frequently expressed by a formula involving algebraic processes or by a trigonometric formula. In such cases we often allow ourselves the convenience of referring to the formula as if it were the function. Thus, for instance, we may speak of the function

$$y = \frac{4x}{x^2 + 4},$$

when what we really mean is "the function consisting of all pairs $(x, y)$, where $y = 4x/(x^2 + 4)$." Of course, when we speak of a function by reference to a formula, it is essential that we understand clearly which letter denotes the independent variable. In most cases the domain of definition of the function is understood to consist of all values of the independent variable for which the formula makes sense. For instance, the function defined by

$$y = \frac{x^2}{x^2 - 1}$$

has for its domain of definition all values of $x$ except $\pm 1$. These two values must be ruled out because a fraction has no meaning if its denominator is 0. *Division by* 0 *is not defined* in algebra.

Functions may be defined without using formulas.

**Example 4:** Let $P$ be the postage (number of cents) required for a letter of $x$ ounces. Then $P$ is a function of $x$. The postal regulations specify postage of 4¢ for each ounce or fraction thereof, up to 70 pounds. Thus $P = 4$ if $0 < x \leq 1$, $P = 8$ if $1 < x \leq 2$, $P = 12$ if $2 < x \leq 3$, and so on. The function is defined for all $x$ such that $0 < x \leq 1120$, and the range of values is the set of numbers $4, 8, 12, \cdots, 4480$.

**Example 5:** Here is another example of a function defined without a formula. It depends on the notion of a prime positive integer. A positive integer greater than 1 is called a *prime* if it cannot be evenly divided by any positive integer except itself and 1. The primes up to 53 are

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53.$$

Now, if $x$ is any number such that $x \geq 1$, let $N$ denote the number of primes not larger than $x$. Then $N$ is a function of $x$. We can make a partial table of the pairs $(x, N)$ as follows:

| *Values of $x$* | *Values of $N$* |
|---|---|
| $1 \leq x < 2$ | 0 |
| $2 \leq x < 3$ | 1 |
| $3 \leq x < 5$ | 2 |
| $5 \leq x < 7$ | 3 |
| $7 \leq x < 11$ | 4 |
| $11 \leq x < 13$ | 5 |
| ... | ... |

An equation involving $x$ and $y$ may be used to define $y$ as a function of $x$ if the equation is equivalent to an equation which expresses the value of $y$ *uniquely* in terms of the value of $x$. For example, the equation $2x + 3y = 4$ can be solved for $y$: $y = \frac{1}{3}(4 - 2x)$, and the latter equation determines $y$ as a function of $x$. In some cases it may happen, however, that the process of solving for $y$ in terms of $x$ leads to more than one value of $y$. For example, if the equation is $x^2 + y^2 - 4 = 0$, we get $y^2 = 4 - x^2$, $y = \pm\sqrt{4 - x^2}$. Since this gives us *two* values of $y$ for each $x$ such that $-2 < x < 2$, the equation $x^2 + y^2 - 4 = 0$ does not by itself determine $y$ as a function of $x$. We emphasize that in our definition of a function we specified that there must be a *uniquely determined value* of $y$ corresponding to each allowable value of $x$. If we wish, we can split the formula $y = \pm\sqrt{4 - x^2}$ into two separate formulas, $y = \sqrt{4 - x^2}$ and $y = -\sqrt{4 - x^2}$. Each of these formulas determines $y$ as a function of $x$.

There is a concept of what is called a *multiple-valued function*, but we shall not have much use for this concept. According to this terminology the formula $y = \pm\sqrt{4 - x^2}$ would determine $y$ as a multiple-valued (generally two-valued) function of $x$. Another example of a multiple-valued function is furnished by defining $y$ to be any angle (in degrees) whose sine is the number $x$. Here there can be an infinite number of $y$'s corresponding to a single $x$. For instance,

$$\tfrac{1}{2} = \sin 30° = \sin 150° = \sin 390° = \sin 510° = \cdots.$$

If $x = \frac{1}{2}$, the corresponding values of $y$ are

$$30° + 360n° \quad \text{and} \quad 150° + 360n°, \quad n = 0, \pm 1, \pm 2, \cdots.$$

Hereafter all functions will be understood to be *single-valued* (a unique $y$ for each $x$) unless something specifically to the contrary is indicated.

In certain problems the discovery of the functional relation between the dependent variable and the independent variable involves the study of a geometrical situation.

**Example 6:** A rectangle is fitted inside of an isosceles triangle of base 4 inches and height 6 inches, as shown in Fig. 1-24. Express the area $A$ of the rectangle as a function of its height $y$.

Letter the figure as shown, with $O$ the mid-point of the base. Then $OC = 2$, $OE = 6$, $BD = y$. We denote $OB$ by $x$. By similar triangles we see that

$$\frac{BD}{BC} = \frac{OE}{OC},$$

or

$$\frac{y}{2 - x} = \frac{6}{2} = 3.$$

Hence $y = 6 - 3x$. Now the area of the rectangle is

$$A = 2xy.$$

We want $A$ expressed in terms of $y$, so we eliminate $x$: $x = \frac{1}{3}(6 - y)$. Then

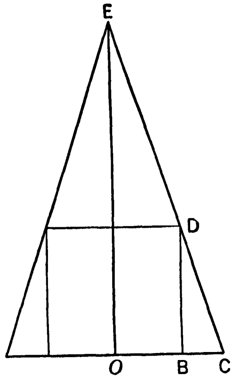


**Fig. 1-24**

$$A = \frac{2}{3}(6y - y^2)$$

is the required formula expressing $A$ as a function of $y$.

## *Graphical Representation of Functions*

If $y$ is a function of $x$, we can represent the function graphically, using rectangular coordinates. The standard method is to use the independent variable as the abscissa and the dependent variable as the ordinate. Thus each number pair $(x, y)$ belonging to the function corresponds to a point in the $xy$-plane, and the collection of all these points makes a configuration called the *graph of the function*. The graph is a visual aid, helping us to comprehend the nature of the function. It shows us at a glance many things about the function, for instance: how changes in $x$ affect the value of $y$ (e.g., whether $y$ increases or decreases as $x$ increases from some specified value), whether there are abrupt changes in $y$ for slight changes in $x$, whether all numbers are possible as values of $y$, and if not, which numbers *do* occur as values of $y$.

The graph of a function defined by a formula is a special case of the graph of an equation, for the formula is an equation.

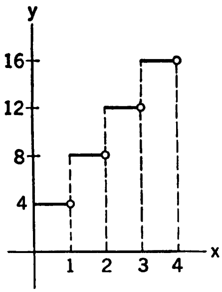



**Fig. 1-25**

The graph of a function need not be a smooth unbroken curve. We illustrate this by showing the graph of the postage function (Example 4). For convenience we use a different (smaller) scale on the $P$-axis from that used on the $x$-axis.* Recall that $P = 4$ if $0 < x \leq 1$, and that in general $P = 4n$ if $n - 1 < x \leq n$ $(n = 1, 2, \cdots, 1120)$. Part of the graph is shown in Fig. 1-25. The tiny circles indicate that the right-hand ends of the horizontal line segments belong to the graph. The left-hand ends of the segments do not belong to the graph. The graph consists solely of these disconnected horizontal segments.

## EXERCISES

1. A rectangle is required to have an area of 4 square feet, but its dimensions may vary. If one side has length $x$, express the perimeter $P$ of the rectangle as a function of $x$.

2. Express the area $A$ of an equilateral triangle as a function of its side, of length $x$.

3. A rectangular pasture, with one side bounded by a straight river, is fenced on the remaining three sides. If the length of the fence is 200 yards, express the area of the pasture as a function of the length of the side along the river.

4. A baseball diamond is a square 90 feet on a side. A player is running from home to first base at the rate of 30 feet per second. Express the runner's distance $s$ from second base as a function of the time ($t$ seconds) since he left home plate.

5. A rectangle of dimensions $2x$ by $2y$ is inscribed in a circle of radius 10. Express $y$ and the area $A$ of the rectangle as functions of $x$.

6. A right circular cone has altitude equal to $r$, the radius of the base. Express the volume $V$ and lateral surface area $S$ of the cone as functions of $r$.

* It is often convenient, in graphing a function, to use different units of length on the two axes. The basic idea of graphing an equation or a function does not require the use of equal units on the two axes. It is only when we discuss certain notions of Euclidean geometry in the plane, e.g., distance between points, or angles between lines, that the use of equal units on both axes is essential.

**7.** A right circular cylinder is inscribed in a sphere of radius 4. Express the volume $V$ and total surface area $S$ of the cylinder as functions of its altitude $y$.

**8.** For each $x$ let $y$ be the nonnegative root of the quadratic equation $y^2 + y - x^2 = 0$. Express $y$ as a function of $x$ by a formula and draw the graph of the function.

**9.** Draw the graph of each of the following functions. The function in (c) is not defined when $x = 0$. In the case of (d) note especially the difference between the graphs of $y = |1 - x^2|$ and $y = 1 - x^2$ when $x^2 > 1$.

(a) $y = |x|$. (c) $y = \dfrac{x}{|x|}$.

(b) $y = x + |x|$. (d) $y = |1 - x^2|$.

**10.** Consider the set of all pairs $(x, y)$, where $x$ can be any number, and $y$ is always 3. Is this a function? What formula expresses the nature of this collection of pairs?

**11.** Is the line $x = 3$ the graph of a function, with $x$ the independent variable?

**12.** (a) For each $x$ let $n$ denote the algebraically largest integer such that $n \leq x$. For instance, $n = -3$ if $x = -\frac{5}{2}$, $n = 0$ if $x = \frac{3}{4}$, $n = 2$ if $x = 2$. Graph the function consisting of all pairs $(x, n)$.
(b) Let $y = x - n$, where $n$ is defined in (a). Draw the graph of the function consisting of all pairs $(x, y)$.

**13.** Consider the square with corners at the points (0, 1), (1, 1), (1, 2), (0, 2). If $P$ is the point $(x, 0)$, let $D$ be the shortest distance from $P$ to a point on the perimeter of the square. Then $D$ is a function of $x$, but to express this function by formulas requires three different formulas, according as $x < 0$, $0 \leq x \leq 1$, or $1 \leq x$. Write out these three formulas. Draw the graph, considering especially values of $x$ such that $-3 \leq x \leq 5$.

**14.** (a) Suppose $a$ and $b$ are positive constants. Let the point $(0, a)$ represent an offshore rock in the ocean. Let the $x$-axis denote the shore line, with $(b, 0)$ a lifeguard station. The mile is the unit of distance. A man swims from the rock to the point $(x, 0)$ and then runs to the lifeguard station. If he swims $s$ miles per hour and runs $r$ miles per hour, and if $T$ is the total time required for the trip, express $T$ as a function of $x$. Consider only values of $x$ such that $0 \leq x \leq b$.
(b) Choose $a = \frac{3}{4}$, $b = 1$, $s = 2$, $r = 6$ and construct a rough graph of the function, using $x = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$.
(c) Proceed as in (b) if $s = 3$, $r = 4$, and $a$ and $b$ are as before.

**15.** The $x$-axis represents the ground level. The line segment from (0, 0) to $(0, b)$ represents a fence $b$ feet high. A ladder 12 feet long has one end at $(a, 0)$, where $a \leq 0$, and the ladder is propped over the fence, with the other end at $(x, y)$ in the first quadrant. (a) Find $x$ as a function of $a$ if $b = 6$ and $-6\sqrt{3} \leq a \leq 0$. (b) Draw the graph of $x$ as a function of $a$. What happens to $x$ as $a$ increases from $-6\sqrt{3}$ to 0?

## 1-7 The Derivative of a Function. Velocity and Acceleration

*Instantaneous Velocity*

As a preliminary to the discussion of the general concept of the derivative of a function, we shall consider the notion of the instantaneous velocity of an object which moves along a straight line in accordance with some definite law. We suppose that a number scale has been established on this line, and we call the line the $s$-axis, using $s$ as the coordinate of the point where the object is located on the axis. We measure time from some chosen instant, letting $t$ be the number of time units; then $t > 0$ for instants after the initial instant and $t < 0$ for instants prior to the initial instant. The moving object, whatever its size or shape, is for present purposes thought of as a single point in motion. A law of motion is simply a definite rule which establishes $s$ as a function of $t$.

**Example 1:** Suppose that $s = 128t - 16t^2$. This formula describes the motion of a ball thrown straight up in the air with a certain initial velocity. It is assumed that the $s$-axis is directed positively upward, with the origin at the point corresponding to $t = 0$, where the ball is thrown. Units of time and distance are seconds and feet.

**Example 2:** Suppose that $s = 200t^2 + 20t + \frac{1}{2}$. This formula might describe, during the limited time $-\frac{1}{20} \leq t \leq \frac{1}{5}$, the motion of a train which is gaining speed at a constant rate. Assume that $t$ is measured in hours and $s$ in miles. The significance of the time interval $-\frac{1}{20} \leq t \leq \frac{1}{5}$ is that the train starts from rest at $t = -\frac{1}{20}$ and reaches a speed of 100 miles per hour at $t = \frac{1}{5}$ (i.e., 15 minutes later). This example is subject to discussion in Exercise 2.

**Example 3:** The formula $s = 400t - 6000t^2 + 30{,}000t^3$ might describe the motion of an inert projectile striking an earthen bank and penetrating a certain distance. If $t$ and $s$ are both 0 at impact and increase as the projectile penetrates, the stated formula gives a reasonable law of motion during the time interval $0 \leq t \leq \frac{1}{15}$. As we shall see presently, the projectile will have stopped when $t = \frac{1}{15}$. Units of time and distance are seconds and feet.

The laws of motion in the foregoing examples are presented without analysis or derivation. We shall examine them in more detail later.

We now ask these questions: What do we mean by the velocity of a moving body at a given instant? How do we find this velocity if the law of motion is known? We are all familiar with velocity or speed in the general sense of a number measuring the rate of traversing distance. We speak of walking 4 miles per hour, of driving 60 miles per hour, and so on. We also speak of *average* speeds, and these are the quantities we are actually accustomed to computing. An average speed is a simple quotient:

$$\text{average speed} = \frac{\text{distance traversed}}{\text{time elapsed}}.$$

It is not so simple, however, to determine the exact velocity of a moving body at a given instant, if the body does not traverse equal distances in equal times. This notion of exact velocity at a given instant, we call it *instantaneous velocity*, is defined by using in an appropriate way the average velocity over shorter and shorter intervals of time.

Consider a definite law of motion, so that $s$ is a function of $t$. Suppose we wish to define the exact velocity for this motion at the instant $t = t_0$ ($t_0$ denoting some fixed value of $t$). Let the corresponding value of $s$ be $s_0$. For any $t$ different from $t_0$, and the corresponding $s$, the quotient

$$\frac{s - s_0}{t - t_0}$$

is called the *average velocity* for the motion during the time interval between the two instants. It does not matter whether $t$ is before or after $t_0$. The average velocity may be either positive, negative, or zero, depending on the particular situation. Now we consider the average velocity for all values of $t$ near $t_0$, and we investigate what happens as $t$ approaches $t_0$. We may expect the average velocity to have different values as we change $t$. *But if it happens that the average velocity approaches a definite limiting value as t approaches $t_0$, this limiting value is defined to be the exact velocity at the instant $t_0$.* We denote it by $v_0$ and indicate the process of getting $v_0$ by writing

$$v_0 = \lim_{t \to t_0} \frac{s - s_0}{t - t_0}. \tag{1}$$

In this process it is of course understood that we confine our attention to the values of $t$ allowed by the law of motion. If values of $t$ on both sides of $t_0$ are permitted, they must be considered.

The units for velocity depend on the units of distance and time, so we have feet per second, centimeters per second, miles per hour, and so on.

Next we shall illustrate this definition of velocity in connection with some of the examples mentioned earlier.

**Example 1, continued:** Find the velocity in the case of the motion $s = 128t - 16t^2$.

Here we have

$$\begin{aligned} s - s_0 &= 128t - 16t^2 - 128t_0 + 16t_0^2 \\ &= 128(t - t_0) - 16(t^2 - t_0^2) \\ &= (t - t_0)[128 - 16(t + t_0)], \end{aligned}$$

$$\frac{s - s_0}{t - t_0} = 128 - 16(t + t_0). \tag{2}$$

At one of the steps we used the formula

$$t^2 - t_0^2 = (t - t_0)(t + t_0).$$

Now, as $t$ gets closer and closer to $t_0$, then $t + t_0$ becomes more and more nearly equal to $2t_0$; therefore we see that in this case

$$v_0 = \lim_{t \to t_0} \frac{s - s_0}{t - t_0} = 128 - 32t_0.$$

If we drop the subscript we have general formula for the velocity $v$ at any instant:

$$v = 128 - 32t. \tag{3}$$

In particular, $v = 128$ if $t = 0$, so the ball was thrown with an initial velocity of 128 feet per second. We see from (3) that the velocity decreases as $t$ increases from 0 to 4 and that $v = 0$ when $t = 4$. This is the instant at which the ball reaches its highest point. After this the ball falls and the velocity becomes negative. As we shall see later, it is a general fact in all motions on the $s$-axis that a positive velocity indicates increasing $s$ with increasing $t$, while negative velocity indicates decreasing $s$ with increasing $t$.

Discussion of the velocity for the motion described in Example 2 is left for the exercises.

**Example 3, continued:** Find the velocity in the case of the motion $s = 400t - 6000t^2 + 30{,}000t^3$.

Here we have

$$s - s_0 = 400(t - t_0) - 6000(t^2 - t_0^2) + 30{,}000(t^3 - t_0^3).$$

By factoring $t - t_0$ out of each term on the right and then dividing through by $t - t_0$ we get

$$\frac{s - s_0}{t - t_0} = 400 - 6000(t + t_0) + 30{,}000(t^2 + tt_0 + t_0^2).$$

As $t$ approaches $t_0$ we see that $t^2$ and $tt_0$ both approach $t_0^2$. Consequently

$$v_0 = \lim_{t \to t_0} \frac{s - s_0}{t - t_0} = 400 - 12{,}000t_0 + 90{,}000t_0^2.$$

Dropping the subscript, we have

$$v = 400 - 12{,}000t + 90{,}000t^2. \tag{4}$$

This can be written

$$v = 90{,}000\left(t - \frac{1}{15}\right)^2. \tag{5}$$

The velocity is 400 feet per second when $t = 0$; it decreases from 400 to 0 as $t$ increases from 0 to $\frac{1}{15}$. It is merely during this short time interval that the formula describing the motion has any physical significance for the projectile. We observe as a matter of interest that $s = \frac{80}{9}$ when $t = \frac{1}{15}$. Thus the projectile penetrates nearly 9 feet before coming to rest.

*The Derivative of a Function*

When a point moves on the $s$-axis, its velocity at a given instant may be described as the rate of change of $s$ with respect to $t$ at that instant.

This concept of rate of change of one quantity with respect to another can be used in many contexts. The pitch of a roof and the steepness of a road up a mountain are examples of rates of change of vertical distance with respect to horizontal distance. In physics, the term *power* means the rate of change of work done with respect to time. In certain kinds of problems the magnitudes of forces are found as rates of change of potential energy with respect to distance.

The general rate of change concept can be considered in the case of any function, if the domain of definition of the function includes some entire interval of the number scale for the independent variable. Suppose $y$ is a function of $x$, and let $x_0$ be a value of $x$ belonging to some interval in the domain of definition of the function. Let $x$ be different from $x_0$. We consider the ratio

$$\frac{y - y_0}{x - x_0},$$

and the limiting value, or limit, of this ratio (if such a limiting value exists) as $x$ approaches $x_0$. *This limiting value is what we define to be the exact rate of change of $y$ with respect to $x$, at $x_0$.* In the standard terminology of calculus this rate of change is called *the derivative of $y$ with respect to $x$*, at $x_0$. One standard notation for this derivative is

$$\left(\frac{dy}{dx}\right)_{x=x_0}.$$

Consequently, by definition,

$$\left(\frac{dy}{dx}\right)_{x=x_0} = \lim_{x \to x_0} \frac{y - y_0}{x - x_0}. \tag{6}$$

In verbal form this definition is: The derivative of $y$ with respect to $x$ at $x_0$ is defined as the limit approached by the quotient $(y - y_0)/(x - x_0)$ as $x$ approaches $x_0$. The assumption is that the quotient *does* approach a limit; in this case we say the derivative *exists*. Otherwise the derivative is not defined. It is furthermore assumed that values of $x$ on both sides of $x_0$ must be considered if the function is defined for such values of $x$; otherwise $x$ is confined to that side of $x_0$ on which the allowable values of $x$ lie. For instance, if $y = x\sqrt{x}$ and $x_0 = 0$, negative values of $x$ are not allowed.

The derivative concept is not used exclusively with the "rate of change" idea foremost. Hence, in the general development of the methods of calculus, we ordinarily use the word "derivative" rather than the "rate of change" terminology.

The notation $dy/dx$ without the parentheses and the suffix $x = x_0$ denotes the value of the derivative of $y$ with respect to $x$ for an arbitrary value of $x$.

**Example 4:** Find the derivative if the function is defined by $y = 3x^4$. We have

$$y - y_0 = 3x^4 - 3x_0^4 = 3(x - x_0)(x^3 + x^2x_0 + xx_0^2 + x_0^3),$$

$$\frac{y - y_0}{x - x_0} = 3(x^3 + x^2x_0 + xx_0^2 + x_0^3).$$

Note the factorization of $x^4 - x_0^4$. As $x$ approaches $x_0$ we obtain

$$\left(\frac{dy}{dx}\right)_{x=x_0} = \lim_{x \to x_0} 3(x^3 + x^2x_0 + xx_0^2 + x_0^3) = 12x_0^3,$$

for $x^3$, $x^2x_0$, and $xx_0^2$ each approaches $x_0^3$. Dropping the subscript, we have the general formula

$$\frac{dy}{dx} = 12x^3.$$

In terms of the derivative terminology and notation we observe that velocity is the derivative of $s$ with respect to $t$, that is,

$$v = \frac{ds}{dt}. \tag{7}$$

*Acceleration*

In studying the motion of an object along a line we are interested, not only in the velocity, but in changes in the velocity. The rate of change of velocity with respect to time is called *acceleration.* If we denote the acceleration at time $t$ by $a$, then

$$a = \frac{dv}{dt}, \tag{8}$$

or, in words, *acceleration is the derivative of velocity with respect to time.* This means, of course, that the acceleration $a_0$ at time $t_0$ is the limit of a certain quotient, namely,

$$a_0 = \lim_{t \to t_0} \frac{v - v_0}{t - t_0}.$$

**Example 1, continued:** Find the acceleration in the case of the motion $s = 128t - 16t^2$.

We know by formula (3) that the velocity is $v = 128 - 32t$. Therefore

$$v - v_0 = 128 - 32t - 128 + 32t_0 = -32(t - t_0),$$

$$\frac{v - v_0}{t - t_0} = -32.$$

In this case the quotient has a constant value, so its limit *is* this value; therefore the acceleration is $a_0 = -32$. This holds for any value of $t_0$, so

$$a = -32.$$

The unit of acceleration is 1 velocity unit per unit of time. In this case the acceleration is $-32$ feet per second per second. The fact that the acceleration is negative indicates that the velocity is decreasing in an algebraic sense. This conforms to our experience in the case of the thrown ball. On the upward flight the velocity is positive and diminishing. On the downward flight the velocity is negative. The ball gathers speed as it falls, but the change of $v$ from 0 to $-32$ to $-64$, and so on, is an algebraic decrease. The fact that the acceleration is constant is the expression of the fundamental law of gravity.

**Example 3, continued:** Find the acceleration in the case of the motion $s = 400t - 6000t^2 + 30{,}000t^3$.

We know by formula (4) that the velocity is $v = 400 - 12{,}000t + 90{,}000t^2$. We leave it for the student to verify in detail the following calculation:

$$\frac{v - v_0}{t - t_0} = -12{,}000 + 90{,}000(t + t_0),$$

$$\left(\frac{dv}{dt}\right)_{t=t_0} = a_0 = -12{,}000 + 180{,}000t_0,$$

$$\frac{dv}{dt} = a = -12{,}000 + 180{,}000t. \tag{9}$$

The significant values of $t$ in this problem are from 0 to $\frac{1}{15}$. During this time the acceleration is negative; it changes from $-12{,}000$ to 0 feet per second per second. The large negative acceleration at $t = 0$ indicates that the velocity, initially 400 feet per second, is decreasing very rapidly. This is the effect of the resistance offered by the earthen bank into which the projectile is going. At $t = \frac{1}{15}$ the velocity and acceleration both reach the value 0. By comparing (9) and (5) it may be noted that

$$a = -600\sqrt{v}.$$

This shows that the magnitude of the acceleration is proportional to the square root of the velocity.

## *The Derivative of a Polynomial*

Expressions such as

$$7, \quad -3 + 2x, \quad 1 - 4x + 5x^2, \quad A + Bx + Cx^2 + Dx^3$$

are called *polynomials* in $x$. An individual term of a polynomial is either a constant or a constant times a power of $x$, the exponent being a positive whole number. A polynomial is an expression which is the sum of a finite number of such terms. Thus the general form of a polynomial is

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

where the coefficients $a_0, a_1, \cdots, a_n$ are constants (i.e., for our present purposes they are real numbers). If $a_n \neq 0$ this polynomial is said to be of degree $n$. Some or all of the coefficients with index less than $n$ may be 0. For instance, we have the following polynomials with degrees as indicated:

$$
\begin{aligned}
7,\ \tfrac{1}{2} &\quad \text{degree 0,}\\
-3 + 2x,\ x &\quad \text{degree 1,}\\
1 - 4x + 5x^2,\ x^2 - 3x,\ 5x^2 &\quad \text{degree 2.}
\end{aligned}
$$

A function defined by setting $y$ equal to a polynomial in $x$ is called a polynomial function; usually we just call it a polynomial, leaving the word "function" to be understood from the context. We shall now state a fundamental theorem about the derivative of a polynomial.

THEOREM 1-B. *Let $y$ be a polynomial in $x$:*

$$y = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n. \tag{10}$$

*Then the derivative of $y$ with respect to $x$ is*

$$\frac{dy}{dx} = a_1 + 2a_2x + \cdots + na_nx^{n-1}. \tag{11}$$

*This result may be stated as follows. The derivative of a polynomial is the sum of the derivatives of its individual terms. The derivative of a constant term is* 0; *the derivative of $cx$ is $c$; and, in general, the derivative of $cx^k$ is $kcx^{k-1}$ (here $c$ denotes a constant coefficient and $k$ is any positive integer).*

*Proof.* If $y$ is given by (10) we have

$$y - y_0 = a_1(x - x_0) + a_2(x^2 - x_0^2) + \cdots + a_n(x^n - x_0^n).$$

We now factor out $x - x_0$ and divide by this factor. The result is

$$\frac{y - y_0}{x - x_0} = a_1 + a_2(x + x_0) + \cdots + a_n(x^{n-1} + x^{n-2}x_0 + \cdots + x_0^{n-1}).$$

When $x$ approaches $x_0$ we see that $x + x_0$ approaches $2x_0$, $x^2 + xx_0 + x_0^2$ approaches $3x_0^2$, and so on. Thus

$$\left(\frac{dy}{dx}\right)_{x=x_0} = a_1 + 2a_2x_0 + 3a_3x_0^2 + \cdots + na_nx_0^{n-1}.$$

On dropping the subscripts we obtain the formula (11).

This proof, and every example in which we have calculated a derivative, has made use of certain facts about limiting values. For instance, we have found the limiting value of a sum by taking the sum of the limiting values of the individual terms in the sum, and we have found the limiting values of products by taking the product of the limiting values of the factors. A more detailed study of the concepts and rules relating to the finding of limits will be made in the following section.

With the result of Theorem 1-B available to us we can in the future write down the derivative of any polynomial at sight. It does not matter, of course, what letters are used for the variables.

**Example 5:** If a motion on the $s$-axis is defined by the formula

$$s = 3t^4 - 28t^3 + 84t^2 - 96t + 25,$$

calculate the velocity and acceleration at any time $t$.

We have

$$v = \frac{ds}{dt} = 12t^3 - 84t^2 + 168t - 96,$$

$$a = \frac{dv}{dt} = 36t^2 - 168t + 168.$$

The derivative symbolism is often used in the following manner: Instead of writing such formulas as

$$y = 3x^4 - 5x^2, \qquad \frac{dy}{dx} = 12x^3 - 10x,$$

we simply write

$$\frac{d}{dx}(3x^4 - 5x^2) = 12x^3 - 10x^2.$$

That is, $\frac{d}{dx}(\quad)$ denotes the derivative with respect to $x$ of whatever is placed inside the parentheses. The object placed inside the parentheses must, of course, define a function of $x$.

## EXERCISES

Use Theorem 1-B in finding derivatives unless the Exercise directs otherwise.

1 For each of the following motions find the velocity $v$ and acceleration $a$.
(a) $s = 96t - \frac{1}{2}t^3$. When is $v$ positive and when negative? Is $v$ increasing or decreasing when $t > 0$?
(b) $s = 256 + 96t - 16t^2$. What is the value of $s$ when $v = 0$? Does $v$ ever increase, algebraically?
(c) $s = t^3 - 9t^2 + 15t - 7$. Find the values of $s$ and $v$ when the acceleration is 0. For what values of $t$ is $v < 0$?
(d) $s = 6t^2 - 2t^3$. By studying the sign of $v$ describe the way $s$ changes when $0 < t < 2$; when $2 < t$. During what interval of time after $t = 0$ is $v$ increasing?
(e) $s = 64t^2 - 16t^4$. By examining the sign of $v$ find the largest positive value that $s$ can attain. What is the acceleration at $t = 0$? at $t = \sqrt{2}$?
(f) $s = 8t^3 - 48t^2 + 72t$. Find the two values of $t$ for which $v = 0$. What values does $s$ have at these two instants? How was $s$ changing before the first of the instants? How does $s$ change between the two instants? What can you say about the increasing or decreasing behavior of $v$ before and after the instant at which the acceleration is 0?

**2.** Consider the motion of Example 2, namely $s = 200t^2 + 20t + \frac{1}{2}$ ($s$ in miles, $t$ in hours). (a) Find the velocity $v$ and the acceleration as functions of $t$. (b) Find the values of $s$ and $t$ when $v = 0$; when $v = 100$. (c) How far does the train go in the first minute after it starts? (d) How far does the train go during the time the velocity increases from 0 to 60?

**3.** If $A$ is the area of a square of side $x$ inches, find the rate of change of $A$ with respect to $x$ when $A$ is 64 square inches.

**4.** (a) Find in terms of $r$ the rate of change of volume $V$ of a sphere with respect to its radius $r$. (b) What is the derivative of $V$ with respect to the diameter $D$ when $D = 6$?

**5.** Find the rate of change of the area of a circle with respect to its diameter when the circumference of the circle is 5 units.

**6.** A conical pile of sand has its height equal to the diameter of its base. As the pile is increased in size, find the rate of increase of its volume with respect to the radius of the base, in terms of this radius.

**7.** A stone dropped into a pool causes a circular ripple to expand, its radius increasing 3 feet per second. How fast is the area within the circle increasing as a function of $t$, the number of seconds after the stone touches the water?

**8.** The point $(x, y)$ moves along the line through $(2, 8)$ and $(6, -2)$. Find the rate of change of $y$ with respect to $x$, and the rate of change of $x$ with respect to $y$.

**9.** Find the rate of change of the area $A$ of an equilateral triangle (a) with respect to its altitude $y$; (b) with respect to the length $x$ of a side.

**10.** A spherical container of radius $R$ feet contains water, the greatest depth being $h$ feet. The volume $V$ of the water is $V = \frac{\pi}{3}h^2(3R - h)$. Find the rate of change of $V$ with respect to $h$ and evaluate at (a) $h = 0$; (b) $h = R$; (c) $h = 2R$.

**11.** Find when $dy/dx = 0$ if
(a) $y = \frac{1}{3}(x^3 - 9x^2 + 15x + 20)$.
(b) $y = 48 + 24x - 3x^2 - x^3$.
(c) $y = 2x^3 - 3x^2 - 12x + 6$.
(d) $y = 21 + 6x - \frac{3}{2}x^2 - x^3$.
(e) $y = x^2 + 2x - \frac{5}{3}$.
(f) $y = x^4 - 2x^3 - 4x^2 + 8x + 1$.

**12.** Find when $ds/dt = 0$ if
(a) $s = t^5 - 30t^3 + 405t$.
(b) $s = 60t^4 + 80t^3 - 450t^2 + 10$.
(c) $s = 3t^5 - 20t^2 + 4$.
(d) $s = t^3 - t^2 - 8t + 6$.
(e) $s = t^4 - 4t^3 + 6t^2 - 4t + 3$.
(f) $s = t^4 - 6t^3 + 12t^2 - 10t + 4$.

13. As an approximate formula, the Fahrenheit temperature $T$ of boiling water is a first-degree polynomial in $h$, where $h$ is the altitude above sea level, in feet. Assuming that $T = 212$ at sea level and $T = 183$ at $h = 14{,}500$, find $dT/dh$ and describe the meaning of its value.

14. For heights up to 500 meters above sea level it is approximately true that the barometric reading $p$ ($p$ millimeters of mercury) is a first-degree polynomial in $h$, where $h$ is altitude above sea level, in meters. If $p = 742$ at $h = 200$ and $p = 715$ at $h = 500$, find $dp/dh$ and describe the meaning of its value in terms of the effect on $p$ of an increase of $h$ by 100 meters.

15. If $F$ and $C$ are corresponding temperatures on the Fahrenheit and centigrade scales (see Exercise 7, § 1-4), find $dF/dC$ and $dC/dF$.

16. A point moves on the $s$-axis so that $s$ is a second-degree polynomial in $t$. If $s = 0$ and $v = 500$ when $t = 0$, and if $dv/dt = -3000$, find $t$ when $v = 0$.

17. An oil tank is being emptied. If there are $G$ gallons of oil in the tank at time $t$, where $G = 67{,}500 - 9000t + 300t^2$ and $t$ is measured in minutes, how many gallons of oil per minute are running out (a) at $t = 0$? (b) one minute before the tank is empty?

18. (a) If $p$ is a function of $q$, write the fraction whose limit as $q$ approaches $q_1$ is, by definition, $(dp/dq)_{q=q_1}$.
(b) If $p = 7q^5$, write out the details of finding $dp/dq$ in a manner analogous to the solution of Example 4.

19. If $y = 1/x$, find $(dy/dx)_{x=2}$ by using the definition in formula (6).

20. If $y = 1/x^2$, find $(dy/dx)_{x=x_0}$ by using the definition in formula (6).

## 1-8 Functional Notation. Limits. Continuity

*Functional Notation*

Suppose $y$ is a function of $x$. For many purposes it is convenient to use a letter to represent the function. If $f$ is the letter selected, the value of $y$ corresponding to $x$ is denoted by $f(x)$, so that $y = f(x)$. Thus $f(2)$ is the value corresponding to $x = 2$. Other illustrations are listed as follows.

| *Value of $x$* | *Corresponding value of $y$* |
|---|---|
| $a$ | $f(a)$ |
| $b$ | $f(b)$ |
| $a + h$ | $f(a + h)$ |
| $\pi$ | $f(\pi)$ |
| $x_0$ | $f(x_0)$ |

In practice we often indicate the definition of a function by setting $f(x)$ equal to an expression involving $x$ (of an algebraic or trigonometric type, for instance), with the understanding that $x$ is the independent variable, and that $x$ may be assigned any value for which the expression makes sense. Other letters may be used, of course, both for the function and for the independent variable.

**Example 1:** $f(x) = \sqrt{1 + x^2}$. Then $f(1) = \sqrt{1 + 1^2} = \sqrt{2}$, $f(3) = \sqrt{10}$, $f(a + h) = \sqrt{1 + (a + h)^2}$.

**Example 2:** $g(t) = \dfrac{t}{1 - t}$. Then $g\left(\dfrac{1}{2}\right) = 1$, $g(-2) = \dfrac{-2}{3}$, $g(t + h) = \dfrac{t + h}{1 - (t + h)}$ (provided that $t + h \neq 1$).

**Example 3:** $F(u) = 2^{-u}$. Then $F(0) = 1$, $F(1) = \frac{1}{2}$, $F(-2) = 4$, $F(x + y) = 2^{-(x+y)} = 2^{-x}2^{-y} = F(x)F(y)$.

**Example 4:** $\phi(x) = \sin x$. Then $\phi(t) = \sin t$, $\phi(0) = 0$, $\phi(x + y) = \sin (x + y) = \sin x \cos y + \cos x \sin y = \phi(x) \cos y + \phi(y) \cos x$.

The definition of the derivative of a function is expressible in terms of functional notation. It is important for the student to be familiar with this. Let us refer back to formula (6) in § 1-7. If $y = f(x)$, we have

$$\frac{y - y_0}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0}.$$

Thus

$$\left(\frac{dy}{dx}\right)_{x=x_0} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{1}$$

The derivative of the function $f$ with respect to $x$ at $x_0$ is often denoted by $f'(x_0)$. (This is read "$f$ prime of $x_0$.") Hence we can write

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{2}$$

In general, the value of the derivative at $x$ is denoted by $f'(x)$.

**Example 5:** If $f(x) = x^5 - 3x^2 + x - 7$, we know by Theorem 1-B that $f'(x) = 5x^4 - 6x + 1$.

If we think of $x$ as a fixed number, and consider a nearby value which is made to approach $x$, the nearby value may be represented by $x + h$, where $h$ is made to approach 0 ($h$ may be either positive or negative). If we use this notation, $f'(x)$ is the limit of the quotient

$$\frac{f(x + h) - f(x)}{(x + h) - x}$$

as $h$ approaches 0, so that

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}. \tag{3}$$

The student should know the definition of the derivative in the various forms which have been given: formula (6) in § 1-7 and the verbal rendering which follows it; and formulas (1), (2), (3) of the present section. The student should recognize the *substance* of these formulas, and be able to write or identify them when other letters are used for the function and the variables.

If $y = f(x)$ and if the derivative of $y$ with respect to $x$ exists for a certain set of values of $x$, the collection of pairs $[x, f'(x)]$ corresponding to these values of $x$ is a function. This function is denoted by $f'$ (read "$f$ prime"), and it is called the derivative of $f$ (with respect to its independent variable). If $f$ has a derivative at $x$, the function is said to be *differentiable* at $x$. The process of finding the derivative of a function is called *differentiation.*

*Limiting Values of Functions*

In our definition of a derivative we use the concept of the limiting value of a quotient; see (2) or (3). In the particular cases which we have considered in § 1-7 the quotient was first simplified by algebra and then we had to find the limiting value of the resulting expression. For instance, in Example 4 of § 1-7 we made the assertion that

$$\lim_{x \to x_0} 3(x^3 + x^2x_0 + xx_0^2 + x_0^3) = 12x_0^3.$$

Now we shall define and explain the meaning of a statement of the following general form: "$f(x)$ approaches the number $A$ as a limit when $x$ approaches $x_0$." An abbreviated symbolic form of this statement is $f(x) \to A$ as $x \to x_0$. We also say "the limit of $f(x)$ is $A$ as $x$ approaches $x_0$." Another way of writing it in symbols is

$$\lim_{x \to x_0} f(x) = A.$$

**Example 6:**

(a) $x^2 + 2 \to 11$ as $x \to 3$.

(b) $\dfrac{x - 2}{x + 4} \to -5$ as $x \to -3$.

(c) $(x^2 - 1)(2x^3 - x) \to 42$ as $x \to 2$.

As a first attempt to make a general definition of what it means to say that $f(x) \to A$ as $x \to x_0$, let us express it this way: We consider the values of $f(x)$ for $x$ near, but not equal to, $x_0$. We examine the effect on $f(x)$ of bringing $x$ closer and closer to $x_0$. If the effect is to bring $f(x)$ closer and closer to the number $A$, and if we can bring the absolute value $|f(x) - A|$ down to any desired smallness and maintain it that small, or even smaller, simply by insisting on an adequate smallness of the absolute value $|x - x_0|$, then we say that $f(x)$ approaches $A$ as $x$ approaches $x_0$. In this definition we *do not* insist that $f(x)$ approach $A$ *without reaching* $A$. It might well happen that $f(x)$ oscillates back and forth, passing through the value $A$ an infinite number of times, but with diminishing amplitude of oscillation, as $x$ moves steadily toward $x_0$. Or the behavior may be even more complicated. The important thing is that we can control the size of $|f(x) - A|$ down to any desired smallness by a suitable control on the smallness of $|x - x_0|$, assuming all the while that $x \neq x_0$ and that $x$ stays within the domain of definition of the function.

**Example 7:** Let us see how the "control of smallness" works in the case of the limit assertion $\lim_{x \to 3} (x^2 + 2) = 11$. Here $f(x) = x^2 + 2$ and $A = 11$, so $f(x) - 11 = x^2 - 9 = (x + 3)(x - 3)$. Since we are considering values of $x$ near 3, we may safely limit our attention to values of $x$ between 2 and 4. Then $x + 3$ is between 5 and 7, and hence certainly

$$|f(x) - 11| = |(x + 3)(x - 3)| \leq 7|x - 3|.$$

It is now evident that $|f(x) - 11|$ is not more than 7 times as large as $|x - 3|$, so if we want a certain smallness for $|f(x) - 11|$ we can attain it by insisting that $|x - 3|$ be not over one seventh of the size allowed for $|f(x) - 11|$. For instance,

$$|f(x) - 11| < \frac{1}{100} \quad \text{if} \quad |x - 3| < \frac{1}{700},$$

$$|f(x) - 11| < \frac{1}{1000} \quad \text{if} \quad |x - 3| < \frac{1}{7000},$$

and in general, if $k$ is any positive number,

$$|f(x) - 11| < k \quad \text{if} \quad |x - 3| < \frac{k}{7}.$$

In this last case there is also the proviso that $2 \leq x \leq 4$. This proviso is automatically attended to if $k \leq 7$.

The definition of a limit for $f(x)$ can be stated very concisely by use of inequalities, as follows: We write $\lim_{x \to x_0} f(x) = A$ and say that $f(x) \to A$ as $x \to x_0$ if corresponding to each positive number $k$ there is some positive number $h$ such that, for $x$ in the domain of definition of $f$, it is true that

$$|f(x) - A| < k \quad \text{whenever} \quad 0 < |x - x_0| < h.$$

This form of the definition is logically complete in itself. It expresses briefly all that has previously been expressed in the discussion, and it is the basis for logical reasoning stemming from the limit concept.

*The Basic Theorems Used in Finding Limits*

In actual practice we do not always deal with limits by working through the details of inequalities as was done in Example 7. There are certain theorems about limits which are of great convenience. In most elementary work we use these theorems as formal rules of reckoning far more than we use the actual definition of a limit by inequalities. The three theorems of greatest usefulness deal respectively with sums, products, and quotients of functions. We suppose in stating these theorems that $f(x)$ and $g(x)$ are defined for all values of $x$ in some interval containing $x_0$, except possibly at $x_0$ itself.

THEOREM 1-C. *If $f(x) \to A$ and $g(x) \to B$ as $x \to x_0$, then $f(x) + g(x) \to A + B$ as $x \to x_0$.*

This theorem is sometimes stated in the form: *the limit of a sum is the sum of the limits.* The theorem is extended by mathematical induction to the case of $n$ functions, where $n$ is any positive integer such that $n \geq 2$.

COROLLARY. *If $\lim_{x \to x_0} f_i(x) = A_i$ for $i = 1, 2, \cdots, n$, then*

$$\lim_{x \to x_0} [f_1(x) + \cdots + f_n(x)] = A_1 + \cdots + A_n.$$

THEOREM 1-D. *If $f(x) \to A$ and $g(x) \to B$ as $x \to x_0$ then $f(x)g(x) \to AB$ as $x \to x_0$.*

The brief verbal form of this theorem is: *the limit of a product is the product of the limits.*

COROLLARY. *If $\lim_{x \to x_0} f_i(x) = A_i$ for $i = 1, 2, \cdots, n$, then*

$$\lim_{x \to x_0} [f_1(x)f_2(x) \cdots f_n(x)] = A_1A_2 \cdots A_n.$$

THEOREM 1-E. *If $f(x) \to A$ and $g(x) \to B$ as $x \to x_0$, and if $B \neq 0$, then $\dfrac{f(x)}{g(x)} \to \dfrac{A}{B}$ as $x \to x_0$.*

This theorem has an important condition not needed in the others, namely, that $B \neq 0$. *The limit of a quotient is the quotient of the limits, provided the limit of the denominator is not zero.*

If $B = 0$ we cannot conclude anything certain about the limit of the quotient. This is because a fraction with denominator 0 is meaningless.

The proofs of the foregoing three theorems are made by using the formal definition of a limit. Nothing more is required for the proofs than an understanding of the definition and some ability in reasoning with inequalities. Just at present we attach more importance to becoming aware of the theorems and their uses than to the giving of the proofs, so it will be our policy to proceed with the development of calculus, using the theorems freely. The proofs are given in § 14-2. In a rough intuitive sense the theorems are "obviously true." That is, it seems apparent (from experience with arithmetic) that if $f(x)$ is near $A$ and $g(x)$ is near $B$, then $f(x) + g(x)$ is near $A + B$, $f(x)g(x)$ is near $AB$, and $f(x)/g(x)$ is near $A/B$, provided $B \neq 0$. Statements like this are merely rephrasings of the meanings of the theorems, however; they are not proofs.

**Example 8:** Find $\lim_{x \to 2} \frac{3x^2 - 2x + 7}{x^3 + 5x}$ and point out how Theorems 1-C, 1-D, and 1-E are used in the solution.

First we observe that $3x^2 \to 3 \cdot 2 \cdot 2 = 12$ as $x \to 2$. This is by an application of the corollary of Theorem 1-D, with $f_1(x) = 3, f_2(x) = x, f_3(x) = x$. The product rule shows likewise that $-2x \to -2 \cdot 2 = -4$, $x^3 \to 2 \cdot 2 \cdot 2 = 8$, and $5x \to 5 \cdot 2 = 10$ as $x \to 2$. The rule for sums (Theorem 1-C and its corollary) then shows that

$$3x^2 - 2x + 7 \to 12 - 4 + 7 = 15$$

and

$$x^3 + 5x \to 8 + 10 = 18$$

as $x \to 2$. Then, by the theorem for quotients,

$$\lim_{x \to 2} \frac{3x^2 - 2x + 7}{x^3 + 5x} = \frac{15}{18} = \frac{5}{6}.$$

A function which is defined by the quotient of two polynomials in $x$ is called a *rational* function of $x$. The function in Example 8 is rational.

If the numerator and denominator of a rational function both happen to be 0 when $x = x_0$, this indicates that the numerator and denominator are both divisible by some power of $x - x_0$. Before attempting to find the limit of the function as $x$ approaches $x_0$, the highest common power of $x - x_0$ should be canceled from numerator and denominator. This is illustrated in the next example.

**Example 9:** Find $\lim_{x \to 3} \frac{x^3 - x^2 - 9x + 9}{x^2 - x - 6}$. As long as $x$ is neither 3 nor $-2$ we have

$$\frac{x^3 - x^2 - 9x + 9}{x^2 - x - 6} = \frac{(x - 3)(x + 3)(x - 1)}{(x - 3)(x + 2)} = \frac{(x + 3)(x - 1)}{x + 2}.$$

Since we require $x \neq 3$ in considering the limit as $x \to 3$, we can write

$$\lim_{x \to 3} \frac{x^3 - x^2 - 9x + 9}{x^2 - x - 6} = \lim_{x \to 3} \frac{(x + 3)(x - 1)}{x + 2} = \frac{12}{5}.$$

### *Continuity*

The noun *continuity* and the adjective *continuous* are used in a special technical sense in mathematics to describe a certain quality which a function may or may not possess at a particular value of $x$. The use of the word "continuous" for this quality is suggested by the everyday use of the word "continuous" to mean "unbroken," or "without interruption." Before giving the exact mathematical definition of continuity, consider an example which illustrates *discontinuity* (i.e., lack of continuity). The "postage function" of Example 4, § 1-6 is discontinuous at each of the points $x = 1, 2, 3, \cdots$, but continuous at all other points $x$ for which it is defined. The discontinuity is expressed by the sudden jump in the value of the function as $x$ passes one of the values $1, 2, 3, \cdots$. On the other hand, the function

$f(x) = x^3$ is continuous for every value of $x$; the continuity is expressed by the fact that small changes in $x$ produce small changes in $x^3$.

The general definition of continuity is this: Suppose a function $f$ is defined at all points of an interval containing the point $x_0$. Then $f$ is said to be continuous at $x_0$ if $f(x)$ approaches a limit as $x$ approaches $x_0$ and if this limit is the value $f(x_0)$ which the function is defined to have at $x_0$. In symbols, the function is continuous at $x_0$ if $\lim_{x \to x_0} f(x) = f(x_0)$.

In elementary calculus it is exceptional for a function to be discontinuous at a point where it is defined. The reason for this is that the very simplest functions which we deal with are continuous at all points, and that most of the processes we use for constructing more complicated functions out of the simple functions retain the continuity of the original constituents. For example, squaring the value of a continuous function gives us a new function which is continuous. Adding or multiplying the values of two continuous functions yields a new function which is continuous, and $f(x)/g(x)$ is continuous at $x_0$ if $f$ and $g$ are continuous at $x_0$ and if $g(x_0) \neq 0$. These assertions follow from the definition of continuity and the Theorems 1-C, 1-D, 1-E about limits.

In particular, a polynomial in $x$ is continuous for all values of $x$, and a rational function of $x$ is continuous for all values of $x$ except those which cause the denominator to equal 0.

A function may be continuous without being differentiable. That is, the mere fact that $f$ is continuous at $x$ does not imply that the derivative $f'(x)$ exists; that is, the quotient $\dfrac{f(x+h) - f(x)}{h}$ may not approach any limit as $h \to 0$. For an illustration see Fig. 1-27 and the remarks pertaining to it. But differentiability *does* imply continuity.

THEOREM 1-F. *If a function is differentiable at a particular point, it is continuous at that point.*

*Proof.* Suppose $f$ is differentiable at $x_0$. We have to prove that $f(x) \to f(x_0)$ as $x \to x_0$. Now, if $x \neq x_0$ we can write

$$f(x) = \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) + f(x_0).$$

Using the rules for limits of products and sums we see that

$$\lim_{x \to x_0} f(x) = f'(x_0) \cdot 0 + f(x_0) = f(x_0).$$

This finishes the proof.

When we know that a function is continuous for all values of $x$ on a certain interval, this helps us to draw the graph of the function. The continuity implies that the part of the graph corresponding to values of $x$

on the interval is without break or interruption. Consequently, what we do in actual practice is to plot a certain number of points of the graph and then join them by an unbroken curve drawn freehand. In doing this we ordinarily use additional information acquired by a study of the function and its derivative.

*Algebra Review*

For use in connection with the exercises at the end of this section, we offer some comments and illustrative examples on the subject of fractions.

To add or subtract fractions, find the least common denominator of the fractions in question. Then express each fraction as an equivalent fraction with the common denominator as its new denominator. After that the addition or subtraction is performed on the new numerators to give the numerator of the result. The denominator of the result is the common denominator.

**Example 10:**

$$\frac{2}{x-2}-\frac{x}{2x+3}=\frac{2(2x+3)}{(x-2)(2x+3)}-\frac{x(x-2)}{(x-2)(2x+3)}$$
$$=\frac{2(2x+3)-x(x-2)}{(x-2)(2x+3)}=\frac{-x^2+6x+6}{(x-2)(2x+3)}.$$

We usually shorten the work by going directly to the single fraction with the common denominator as its denominator.

**Example 11:**

$$\frac{x^2}{x+1}-\frac{3}{x^2-1}+\frac{4x-1}{x+2}=\frac{x^2(x-1)(x+2)-3(x+2)+(4x-1)(x^2-1)}{(x+1)(x-1)(x+2)}$$
$$=\frac{x^4+x^3-2x^2-3x-6+4x^3-x^2-4x+1}{(x+1)(x-1)(x+2)}$$
$$=\frac{x^4+5x^3-3x^2-7x-5}{(x^2-1)(x+2)}.$$

To multiply two fractions we multiply numerators to obtain the numerator of the answer. Likewise for denominators. For instance $\frac{2}{3}\cdot\frac{4}{5}=\frac{8}{15}$. The procedure also applies to algebraic expressions.

To divide by a fraction, we multiply by the inverted fraction. For instance, $\frac{3}{4}\div\frac{7}{8}=\frac{3}{4}\cdot\frac{8}{7}=\frac{6}{7}$. The procedure applies also to algebraic expressions.

**Example 12:**

$$\frac{x^2}{x+1}\div\frac{2x+3}{x^2-4}=\frac{x^2}{x+1}\cdot\frac{x^2-4}{2x+3}=\frac{x^2(x^2-4)}{(x+1)(2x+3)}.$$

We may multiply out if we like in this answer, but in calculus it is often desirable to leave expressions like this in factored form.

Sometimes we have fractions whose numerators and denominators are themselves expressions which involve fractions. In such cases we may begin by simplifying the numerator and denominator separately.

**Example 13:** Simplify the compound fraction

$$\frac{\dfrac{1}{x^2} - 16}{\dfrac{2}{x^3} - \dfrac{3}{x^2} - \dfrac{8}{x}}.$$

We have
$$\frac{1}{x^2} - 16 = \frac{1 - 16x^2}{x^2},$$
$$\frac{2}{x^3} - \frac{3}{x^2} - \frac{8}{x} = \frac{2 - 3x - 8x^2}{x^3}.$$

The original fraction then becomes

$$\frac{1 - 16x^2}{x^2} \cdot \frac{x^3}{2 - 3x - 8x^2} = \frac{(1 - 16x^2)x}{2 - 3x - 8x^2}.$$

Another type of problem which we meet in calculus is illustrated next.

**Example 14:** If $f(x) = \dfrac{1}{3x - 5}$, find

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

We have

$$f(x + h) - f(x) = \frac{1}{3(x + h) - 5} - \frac{1}{3x - 5}$$
$$= \frac{3x - 5 - 3(x + h) + 5}{[3(x + h) - 5](3x - 5)}.$$

The numerator here is

$$3x - 5 - 3x - 3h + 5 = -3h.$$

Thus
$$\frac{f(x + h) - f(x)}{h} = \frac{-3h}{h[3(x + h) - 5](3x - 5)}.$$

We cancel the $h$ and then let $h \to 0$. The result is

$$f'(x) = \frac{-3}{(3x - 5)^2}.$$

## EXERCISES

**1.** If $f(x) = \dfrac{3}{x} - x^3$, find $f(3), f(-1), f\left(\dfrac{p}{q}\right), f\left(\dfrac{1}{x}\right)$.

**2.** If $f(x) = \dfrac{x^3 + 32}{x + 4}$, find $f(0), f(4), f(q^3), f(x_1), f\left(\dfrac{4}{x}\right)$.

**3.** (1) In each case state the values of $x$ for which $F(x)$ is *not* defined. (2) Then compute successively $F(1)$, $F(-4)$, $F(p)$, $F(a^2)$, $F\left(\frac{1}{x}\right)$, $F(u+v)$.

(a) $F(x) = x^3 - \dfrac{x^5}{10}$.

(b) $F(x) = 8x - \dfrac{64}{x^2}$.

(c) $F(x) = \dfrac{x+1}{x(x+8)}$.

(d) $F(x) = \dfrac{x-2}{x^2-9x+20}$.

(e) $F(x) = \dfrac{\sqrt{x^2-1}}{x+1}$.

(f) $F(x) = \sqrt{\dfrac{x+2}{x}}$.

(g) $F(x) = \dfrac{x^3+27}{x^2-9}$.

(h) $F(x) = \dfrac{x^2-25}{x^3+x^2-6x}$.

**4.** (1) Form the expression $\dfrac{F(x+h)-F(x)}{h}$ and simplify it. (2) Then find the limit as $h \to 0$, and write the result as a formula giving the value of $F'(x)$.

(a) $F(x) = \dfrac{1}{x}$.

(b) $F(x) = \dfrac{5}{x^2}$.

(c) $F(x) = \dfrac{1}{2x+3}$.

(d) $F(x) = \dfrac{1}{x^3}$.

(e) $F(x) = \dfrac{x}{x-1}$.

(f) $F(x) = \dfrac{x^2}{3x-4}$.

**5.** Follow the same directions as in Exercise 4.

(a) $F(x) = \dfrac{1}{2-x}$.

(b) $F(x) = \dfrac{3}{x^4}$.

(c) $F(x) = \dfrac{1-x}{1+x}$.

(d) $F(x) = \dfrac{x}{x^2+1}$.

(e) $F(x) = \dfrac{1}{x(x-1)}$.

(f) $F(x) = \dfrac{x^2}{x+2}$.

**6.** In each of the following the indicated limit is a derivative. Of what function and at what point? Given an answer of the type "The limit is the derivative of the function $f(x) = \cdots$ at the point $x = \cdots$."

(a) $\lim\limits_{x\to c} \dfrac{x^{17}-c^{17}}{x-c}$.

(b) $\lim\limits_{h\to 0} \dfrac{(x+h)^5-x^5}{h}$.

(c) $\lim\limits_{x\to x_0} \dfrac{(1/x)-(1/x_0)}{x-x_0}$.

(d) $\lim\limits_{x\to a} \dfrac{(2/x^2)-(2/a^2)}{x-a}$.

(e) $\lim\limits_{x\to 2} \dfrac{\log_{10} x - \log_{10} 2}{x-2}$.

(f) $\lim\limits_{h\to 0} \dfrac{\sqrt{a+h}-\sqrt{a}}{h}$.

(g) $\lim\limits_{p\to 0} \dfrac{2^p-1}{p}$.

(h) $\lim\limits_{y\to x} \dfrac{F(y)-F(x)}{y-x}$.

**7.** (a) If $f(x) = 10^x$, what relation is there between $[f(x)]^n$ and $f(nx)$? (b) What relation is there between $f(x+y)$, $f(x)$, and $f(y)$?

**8.** (a) If $f(x) = \log_{10} x$, what relation is there between $f(xy)$, $f(x)$, and $f(y)$?
(b) Express $f(x^n)$ in terms of $f(x)$.

**9.** If $f(x) = \log_{10} x$ and $g(x) = 10^x$, what are $f[g(x)]$ and $g[f(x)]$?

**10.** Find each of the limits indicated.

(a) $\lim_{x \to 1} \frac{x^2 - 1}{x^2 - 2x + 1}$.

(b) $\lim_{x \to 2} \frac{x^2 - 5x + 6}{x^2 - 6x + 8}$.

(c) $\lim_{x \to 7} \frac{x^2 - 4x - 21}{x^2 + 2x - 63}$.

(d) $\lim_{x \to 4} \frac{x^3 - 64}{x - 4}$.

(e) $\lim_{x \to x_0} \frac{(1/x^2) - (1/x_0^2)}{x - x_0}$.

(f) $\lim_{x \to -2} \frac{(x + 2)(x^2 - x + 3)}{x^2 + 3x + 2}$.

(g) $\lim_{x \to 1} \frac{x^3 - 3x^2 + 2}{x^3 - 2x + 1}$.

(h) $\lim_{x \to -2} \frac{(3x + 1)(x + 2)^2}{(x^2 - 4)(x^2 + 3x + 2)}$.

**11.** Find each of the limits indicated. Assume $a \neq 0$.

(a) $\lim_{x \to -1} \frac{x^2 - 1}{x^2 - 2x - 3}$.

(b) $\lim_{x \to 1} \frac{x^3 + x^2 - 2x}{x^3 - 2x^2 - x + 2}$.

(c) $\lim_{x \to a} \frac{x^2 - a^2}{x^3 - a^3}$.

(d) $\lim_{x \to -a} \frac{x^3 + a^3}{x^4 - a^4}$.

(e) $\lim_{x \to 1} \frac{(x^2 + 2x - 3)^2}{(x - 1)(x + 5)^2}$.

(f) $\lim_{x \to a} \frac{x - a}{x^2 - a^2}$.

(g) $\lim_{x \to a} \frac{\sqrt{x} - \sqrt{a}}{x - a}$, $(a > 0)$.

(h) $\lim_{x \to a^2} \frac{\sqrt{x} - |a|}{x - a^2}$.

**12.** (a) What is $\lim_{x \to 4} (3x^2 - x + 5)$? (b) If $3 \leq x \leq 5$, show that

$$|(3x^2 - x + 5) - 49| \leq 26|x - 4|,$$

and hence that $|(3x^2 - x + 5) - 49| < k$ if $3 \leq x \leq 5$ and $|x - 4| < k/26$ (assuming $k > 0$). What does this have to do with part (a)?

**13.** (a) What is $\lim_{x \to 5} \frac{1}{x}$? (b) If $4 \leq x \leq 6$ and $|x - 5| < 20k$ (where $k > 0$), show that $\left|\frac{1}{x} - \frac{1}{5}\right| < k$. What does this have to do with part (a)?

**14.** (a) After solving Exercises 12 and 13, find $h$ in terms of $k$, if $k > 0$, so that $\left|\frac{x^2 + 1}{x} - \frac{5}{2}\right| < k$ if $1 \leq x \leq 3$ and $|x - 2| < h$. (b) What statement about limits does this prove?

**15.** (a) Draw a figure showing the graph of a function $f$ such that for $x < 2$ the graph is part of a straight line, for $x > 2$ the same is true (though not the same straight line as when $x < 2$), and $f(2) = 1$, $\lim_{x \to 2} f(x) = 0$. Is $f$ continuous? (b) The same problem as (a), except that $f(2) = -1$, $\lim_{x \to 2} f(x) = -1$.

**16.** Draw a figure showing the graph of a function $f$ such that $f(x)$ is constant if $x < 0$, $f(x) = 12 - 2x$ if $x > 4$, $f$ is everywhere continuous, $f(2) = 3$, and the graph of $f$ for $0 < x < 4$ is a straight line segment. What is $f(4)$? What is $f(0)$?

**17.** (a) To show that the function $f(x) = \sqrt{x}$ is continuous at $x = 0$ we must show that $\lim_{x \to 0} \sqrt{x} = \sqrt{0} = 0$. How small must a positive $x$ be to make

$\sqrt{x} < \frac{1}{100}$? $\sqrt{x} < \frac{1}{900}$? $\sqrt{x} < k$, where $k > 0$?

(b) To show that the function $f(x) = \sqrt{x}$ is continuous at $x = c$ (where $c > 0$), we must show that $\lim_{x \to c} \sqrt{x} = \sqrt{c}$. A proof of this by inequalities can be constructed by using the following suggestions in getting started: By algebra $(\sqrt{x} - \sqrt{c})(\sqrt{x} + \sqrt{c}) = x - c$. Thus

$$\sqrt{x} - \sqrt{c} = \frac{x - c}{\sqrt{x} + \sqrt{c}}, \quad \text{and so} \quad |\sqrt{x} - \sqrt{c}| \le \frac{|x - c|}{\sqrt{c}}.$$

Now tell how small $|x - c|$ must be to guarantee $|\sqrt{x} - \sqrt{c}| < k$.

**18.** For this exercise a certain knowledge of exponentials and logarithms is assumed.

Prove that $\lim_{x \to 0} \frac{1}{2^{1/x^2}} = 0$ by showing that if $0 < k < 1$ and

$$0 < |x| < \left(\frac{\log_{10} 2}{\log_{10} (1/k)}\right)^{1/2},$$

then $0 < \frac{1}{2^{1/x^2}} < k$. Start as follows:

$$\frac{1}{2^{1/x^2}} < k \quad \text{is equivalent to} \quad \frac{1}{k} < 2^{1/x^2};$$

$$\frac{1}{k} < 2^{1/x^2} \quad \text{is equivalent to} \quad \log_{10} \frac{1}{k} < \frac{1}{x^2} \log_{10} 2.$$

Now continue. One needs here the two following facts: (1) For positive numbers $a$ and $b$, $a < b$ is equivalent to $\log_{10} a < \log_{10} b$, and (2) if $c > 0$, $x^2 < c$ is equivalent to $|x| < \sqrt{c}$. Note also that $2^{1/x^2}$ is not defined if $x = 0$, and that it is always positive.

## 1-9 Geometrical Meaning of the Derivative

The derivative of a function has a very important meaning in relation to the graph of the function. Suppose that we are considering the graph of $y = f(x)$, where $f$ is a function which is continuous for each value of $x$ on a certain interval. The corresponding part of the graph is then an unbroken curve. The fact of primary importance about the derivative in relation to the curve is this: *If the function has a derivative at $x_0$ and if $y_0 = f(x_0)$, then the straight line through the point $(x_0, y_0)$ with slope $f'(x_0)$ is tangent to the graph at the point. Conversely, if there is a line tangent to the graph at $(x_0, y_0)$ and if this line is not parallel to the y-axis, then the function has a derivative at $x_0$, the value of the derivative being the slope of the line.*

In order to see the truth of the foregoing assertions we must first be clear about what it means to say that a line is tangent to the graph at a certain point. Let $P_0$ be a point on the curve and let $P$ be a distinct point nearby on the curve. Draw the complete line $L$ through $P_0$ and $P$, and consider how this line varies as $P$ approaches $P_0$. If there is a fixed line $T$ through $P_0$ such that the angle $\alpha$ between $L$ and $T$ approaches 0 as $P$ approaches $P_0$ along the curve, the line $T$ is called the tangent to the curve at $P_0$ (see Fig. 1-26). If the curve extends on both sides of $P_0$, the line $L$ must approach coincidence with $T$ as $P$ approaches $P_0$ from either side.

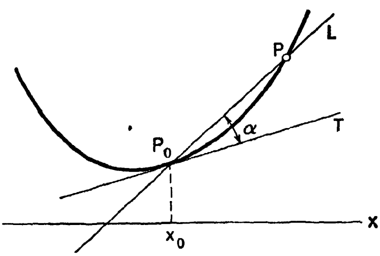



**Fig. 1-26**

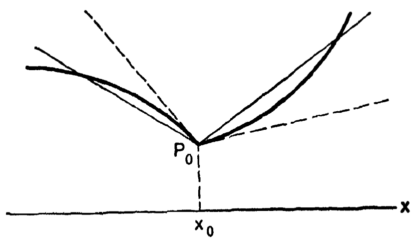



**Fig. 1-27**

If there is no line $T$ which fulfills the foregoing condition, the graph does not have a tangent at $P_0$. The simplest illustration of how this can happen is afforded by a graph consisting of two parts which meet at $P_0$ in the manner depicted in Fig. 1-27. The line through $P_0$ and $P$ approaches two different limiting positions according as $P$ approaches $P_0$ from one side or the other. In this case there is no tangent to the graph at $P_0$. This is a case in which the function is continuous at $x_0$ but not differentiable there.

Now consider the definition of the derivative. Let $P_0$ and $P$ have coordinates $(x_0, y_0)$ and $(x, y)$, respectively, where $y = f(x)$. To have $P$ distinct from $P_0$ means $x \neq x_0$, and to have $P$ approach $P_0$ along the curve is equivalent to having $x \to x_0$, because the function is continuous. The line $L$ through $P_0$ and $P$ has slope

$$\tan \phi' = \frac{y - y_0}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} \tag{1}$$

(see the angle $\phi'$ marked on Fig. 1-28). The condition for the function $f$ to have a derivative at $x_0$ is that the ratio in (1) approach a definite limit [which is $f'(x_0)$] as $x \to x_0$. The condition that the line $T$ through $P_0$, making an angle $\phi_0$ (where $\phi_0 \neq 90°$) with the positive $x$-axis, be tangent to the graph at $P_0$ is that $\phi' \to \phi_0$ as $x \to x_0$. But $\phi' \to \phi_0$ is equivalent to $\tan \phi' \to \tan \phi_0$. *Hence we have a tangent at $P_0$, not parallel to the $y$-axis, if and only if the derivative $f'(x_0)$ exists, and in that case the derivative is the slope of the tangent line.*

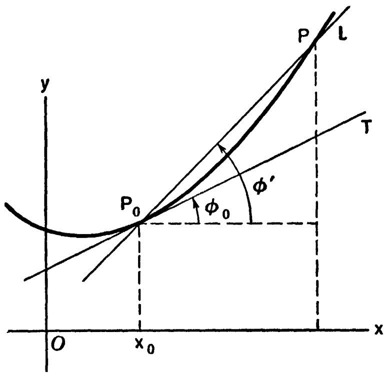



Fig. 1-28

The assertion that $\phi' \to \phi_0$ is equivalent to $\tan \phi' \to \tan \phi_0$ involves two things: (1) that the trigonometric tangent is a continuous function of the angle, and (2) that the angle is a continuous function of its tangent. We take these facts for granted now.

The slope *of a line* was defined in § 1-3. By the slope *of a curve* at a point we mean the slope of the tangent line at that point, if there is a tangent line not parallel to the $y$-axis. By the angle of intersection of two curves we mean the angle of intersection of their tangent lines. Many problems about the tangents to curves can be solved by using the fact that $f'(x_0)$ is the slope of the curve $y = f(x)$ at the point corresponding to $x = x_0$.

**Example 1:** The curves $y = x^2$ and $y = 2 - x^2$ intersect at the points (1, 1) and (−1, 1). Find the equation of the tangents to the curves at (1, 1), and the angle of intersection of the curves at this point (see Fig. 1-29).

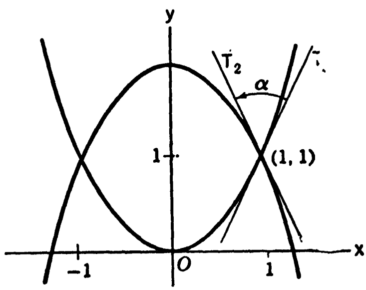



**Fig. 1-29**

The slopes of the curves are given by

$$y = x^2, \qquad \frac{dy}{dx} = 2x = 2 \qquad \text{if} \quad x = 1,$$

$$y = 2 - x^2, \quad \frac{dy}{dx} = -2x = -2 \quad \text{if} \quad x = 1.$$

The tangent $T_1$ to $y = x^2$ at (1, 1) has the equation

$$y - 1 = 2(x - 1), \quad \text{or} \quad 2x - y = 1.$$

The tangent $T_2$ to $y = 2 - x^2$ at (1, 1) has the equation

$$y - 1 = -2(x - 1), \quad \text{or} \quad 2x + y = 3.$$

The angle $\alpha$ from $T_1$ to $T_2$ is determined by the formula

$$\tan \alpha = \frac{-2 - 2}{1 + (-2)\cdot 2} = \frac{-4}{-3} = 1.3333\cdots.$$

A table of tangents shows that $\alpha$ is approximately 53°7′45″.

The line through a point on a curve and perpendicular to the tangent at that point is called the *normal* to the curve at the point. The equation of a normal can be found, in general, by using the fact that the slope of the normal is the negative reciprocal of the slope of the tangent.

If two curves intersect and if at the point of intersection the tangent to one curve is perpendicular to the tangent to the other curve, the two curves are said to intersect *orthogonally*, or to be *orthogonal* at the point of intersection. To find where two curves intersect we solve the two equations as simultaneous equations in $x$ and $y$.

**Example 2:** Find the points of intersection of the curves $y = 1 - \frac{1}{2}x^2$, $y = \frac{1}{4}x^2 - \frac{1}{2}$, and test to see if the curves intersect orthogonally. The curves intersect when

$$1 - \frac{1}{2}x^2 = \frac{1}{4}x^2 - \frac{1}{2}, \quad \text{or} \quad \frac{3}{4}x^2 = \frac{3}{2}.$$

This gives $x = \pm\sqrt{2}$, $y = 0$, so the points of intersection are $(\sqrt{2}, 0)$ and $(-\sqrt{2}, 0)$. The slopes of the curves are respectively $-x$ and $x/2$, so at $(\sqrt{2}, 0)$ they are $-\sqrt{2}$ and $\sqrt{2}/2$. These slopes are negative reciprocals, so the curves intersect orthogonally. The same holds true at the point $(-\sqrt{2}, 0)$.

## EXERCISES

**1.** Find the slope of the curve and the equation of the tangent to the curve at each of the points indicated.
(a) $y = \frac{1}{4}x^3$ at $x = 2$.
(b) $y = \frac{1}{8}x^4$ at $x = \frac{3}{2}$.
(c) $y = 6x^2 - 2x^3$ at $x = -\frac{1}{2}, 0, 1$.
(d) $y = 256 + 96x - 16x^2$ at $x = 0, 3, 8$.
(e) $y = 96x - \frac{1}{2}x^3$ at $x = 0, 8, 14$.

**2.** In each part of this problem proceed as follows: Find the slope of the curve at the indicated points. Plot the points of the curve and use the slopes to draw the tangents at these points. Fill in the curve between the

points indicated, on the assumption that the curve fits smoothly into the framework provided by the tangents.
(a) $y = 2x^2 - 3x + 1$ at $x = 0, \frac{1}{2}, \frac{3}{4}, 1, 2$.
(b) $y = 4 + 4x - 2x^2$ at $x = -1, 0, 1, 2, 3$.
(c) $y = x^3 - 6x^2 + 9x + 1$ at $x = 0, 1, 2, 3, 4$. The curve crosses its tangent at $x = 2$.
(d) $y = 2x^3 - 9x^2 + 12x - 3$ at $x = 0, 1, \frac{3}{2}, 2, 3$. The curve crosses its tangent at $x = \frac{3}{2}$.
(e) $y = \frac{1}{6}x^4 - x^2 + \frac{4}{3}x + \frac{5}{3}$ at $x = -3, -2, -1, 0, 1, 2$. The curve crosses its tangent at $x = -1$ and at $x = 1$.

**3.** (a) Find the tangent to $y = \frac{16}{5}x^2$ at $x = \frac{5}{16}$. (b) Find the two points in which this tangent intersects the curve $y = \frac{1}{4} - x^2$, and the tangents to this latter curve at these points. (c) What angles does the first mentioned tangent make with each of these latter tangents?

**4.** Consider the curve $y = x - \frac{1}{4}x^2$. (a) Find the slope of the curve at the points $x = -1, 0, 2, 4, 5$. Plot the corresponding points on the curve and draw the tangent lines at these points. Sketch in the curve. (b) Now find two points on the curve at each of which the tangent line is such that it goes through the point $(-\frac{1}{2}, 0)$.

**5.** Find the values of $x_0$ such that the tangent to the curve $y = x^3 + x^2 + 4x + 3$ at $x = x_0$ intersects the $x$-axis at $x = \frac{2}{3}x_0$.

**6.** Find the tangent to the curve $4ay = x^2$ at $x = 2a$, where $a > 0$. Prove that the point of tangency, the point where the tangent crosses $y = -2a$, and the origin are vertices of an isosceles triangle. Draw the figure, with $a = 2$.

**7.** Consider the curve $y = \frac{1}{4}x^2$ and the point $F$ with coordinates $(0, 1)$. If $P$ is on the curve in the first quadrant, show that the tangent at $P$ bisects the angle between the line through $F$ and $P$ and the line through $P$ parallel to the $y$-axis. Suggestion: Let $\alpha$ be the angle from $FP$ produced to the tangent at $P$ and let $\beta$ be the angle from the tangent to the line through $P$ parallel to the $y$-axis. Show that $\tan \alpha = \tan \beta$.

**8.** If $x_1 > 0$ find the value of $x_0$ between 0 and $x_1$ such that the tangent to $y = x^3$ at $x = x_0$ is parallel to the chord joining the points corresponding to $x = 0$, $x = x_1$. Draw a figure.

**9.** Consider the curve $y = x^2$ and any two distinct points $(x_1, y_1)$, $(x_2, y_2)$ on it. Find the mid-point of the chord joining these points and let $(x_0, y_0)$ be the point where the line parallel to the $y$-axis and through the mid-point of the chord intersects the curve. Show that the tangent to the curve at $(x_0, y_0)$ is parallel to the chord. Draw a typical figure, say with $x_1 = -1$, $x_2 = 2$. Does the situation work out in the same way for the curve $y = ax^2 + bx + c$, where $a, b, c$ are constants and $a \neq 0$?

**10.** Find the equation of the normal to each curve at the point indicated.
(a) $y = \frac{2}{3}x^2$ at $x = 3$.
(b) $y = \frac{1}{8}x^3$ at $x = 2$.

(c) $y = \frac{9}{16}x^4$ at $x = \frac{4}{3}$.
(d) $y = 2x^3 - 9x^2 + 12x - 3$ at $x = \frac{3}{2}$.
(e) $y = x^3 - 6x^2 + 9x + 1$ at $x = 2$.

11. Find the intersection (or intersections) of each pair of curves, and test to see if the curves intersect orthogonally.
(a) $y = x^2$, $y = x^2 - 2x + 1$.
(b) $y = x^2 - 4x + 4$, $y = x^2 + 2x + 1$.
(c) $y = x^2 - 8x + 16$, $y = x^2 + 6x + 9$.
(d) $y = \frac{1}{2}x^2$, $y = 1 - \frac{1}{2}x^2$.
(e) $y = 2 - x^2$, $y = 2x^2 - 1$.
(f) $y = \frac{1}{2}x^2$, $y = \frac{3}{4} - x^2$.
(g) $y = x^2$, $y = \frac{5}{4} - \frac{1}{4}x^2$.
(h) $y = \frac{1}{4}x^2$, $y = 3 - \frac{1}{8}x^2$.
(i) $y = x^2$, $y = 2 - x^4$.

12. (a) A function $f$ is defined by $f(x) = |x|$. Draw the graph of $y = f(x)$. Consider $x \geq 0$ and $x < 0$ separately. (b) Is there any point of the graph where there is no tangent line? (c) Find $f'(x)$ if $x > 0$; if $x < 0$. What about $x = 0$?

13. Make a diagram showing the graph of a differentiable function $f$ such that $f(-2) = 4$, $f(0) = 2$, $f(2) = 1$, $f(4) = 1$, $f'(-2) = -2$, $f'(0) = 0$, $f'(2) = -1$, $f'(4) = 2$, and there are just two values of $x$ for which $f'(x) = 0$.

14. Draw a graph (not unique) of $y = f(x)$ if $f$ is an everywhere continuous function such that $f'(x)$ is defined if $x < 2$, and $f'(x)$ decreases through positive values, approaching 0 as $x$ increases toward 2, while $f'(x)$ is defined and equal to 1 if $2 < x$. Is $f$ differentiable at $x = 2$?

## 1-10 Increasing and Decreasing Functions

When we were discussing velocity we stated that if $v = ds/dt$ is positive, this implies that $s$ is increasing as $t$ increases. We shall now consider the significance of a positive or negative value of the derivative in the case of any function.

A function $f$ is said to be *increasing* on a certain interval of the $x$-axis if it is defined for each value of $x$ on the interval and if on this interval $x_1 < x_2$ implies $f(x_1) < f(x_2)$. Likewise, the function is said to be *decreasing* on the interval if $x_1 < x_2$ implies $f(x_1) > f(x_2)$.

In drawing the graph of a function it is of great usefulness to know the intervals on which a function is increasing and those on which it is decreasing. This information can be obtained by examining the derivative to see where it is positive and where it is negative. *The function is increasing on any interval throughout which $f'(x) > 0$, and it is decreasing on any interval throughout which $f'(x) < 0$.* The simplest way to justify this assertion is to use a theorem which is proved in the next chapter of the book (the *law of* *the mean,* Theorem 2-C). At our present stage of theoretical development we make the argument somewhat differently. Suppose $f'(x_0) > 0$. Then, since the quotient $[f(x) - f(x_0)]/(x - x_0)$ approaches the positive limit $f'(x_0)$ as $x \to x_0$, the quotient must itself be positive as soon as $x$ is within a certain distance of $x_0$. This means that $f(x) - f(x_0)$ and $x - x_0$ have the same sign. Then, for $x$ in a certain proximity to $x_0$, $f(x) < f(x_0)$ if $x < x_0$, and $f(x) > f(x_0)$ if $x > x_0$. This means that the points of the graph slightly to the left of $(x_0, y_0)$ are lower than this point and those slightly to the right are higher. This kind of argument is valid near each point at which the derivative is positive. Hence, if we start at a point $x_1$ and move to the right on an interval of the $x$-axis where $f'(x) > 0$, we always have $f(x_1) < f(x)$. That is, we never reach a point $x_2$ for which $f(x_2) \leq f(x_1)$. For if we could reach such a point, and if $x_2$ were the *first* such point to the right of $x_1$, the fact that $f'(x_2) > 0$ would imply that for points $x$ near $x_2$ on the left of it we have $f(x) < f(x_2)$, and hence $f(x) < f(x_1)$. This would contradict the fact that $x_2$ is the *first* point to the right of $x_1$ where the function value is less than or equal to $f(x_1)$.

A similar argument shows that if $f'(x) < 0$ on an interval, then the value of $f(x)$ decreases as we move to the right on the interval.

### *Critical Points*

If $f'(x_0) = 0$, the function is said to be *stationary* at $x_0$. The point $x_0$ is called a *critical point* of the function, and the corresponding point of the graph is called a critical point of the graph. A critical point of the graph is recognized by the fact that at this point the slope is 0 and the tangent is parallel to the $x$-axis. Figure 1-30 shows three different critical points.

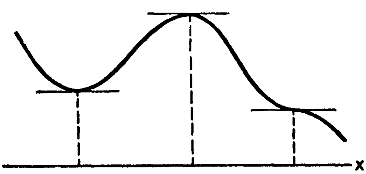



**Fig. 1-30**

If $f(x)$ is a polynomial, the general appearance of the graph of $y = f(x)$ can be determined quite easily once we locate the critical points of the polynomial. In between two critical points the polynomial is either always increasing or always decreasing.

**Example 1:** Construct the graph of $y = 8x^3 - 48x^2 + 72x$. The derivative is

$$\frac{dy}{dx} = 24x^2 - 96x + 72 = 24(x - 1)(x - 3).$$

The critical points are at $x = 1$, $x = 3$. The corresponding values of $y$ are 32 and 0. Now we consider the three possibilities

$$x < 1, \quad 1 < x < 3, \quad 3 < x$$

determined by the position of $x$ in relation to the critical points. When $x < 1$,

$x - 1$ and $x - 3$ are both negative, their product is positive, and so $dy/dx > 0$. If $1 < x < 3$, $x - 1$ is positive and $x - 3$ is negative, so the product is negative and $dy/dx < 0$. If $3 < x$, $x - 1$ and $x - 3$ are both positive and so is $dy/dx$. These results are conveniently displayed on a diagram as in Fig. 1-31. Now

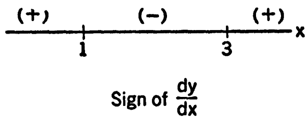
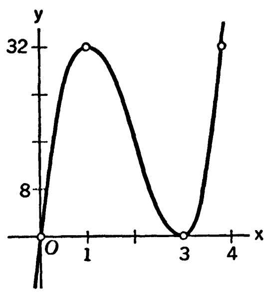



**Fig. 1-31**

we plot the points (1, 32) and (3, 0) and use the fact that $f(x)$ is increasing when $x < 1$, decreasing when $1 < x < 3$, and increasing when $3 < x$. With these items of information we can sketch the general course of the graph. It is helpful to locate at least two more points, one to the left of $x = 1$ and one to the right of $x = 3$. We use $x = 0$ and $x = 4$. The graph is shown in Fig. 1-32. We needed only four points to draw it. These points appear in the table of values. Observe that we have used different scales on the two axes.




| $x$ | $y$ |
|---|---|
| 0 | 0 |
| 1 | 32 |
| 3 | 0 |
| 4 | 32 |

**Fig. 1-32**

### *Some Facts About Polynomials*

We summarize here some important matters of algebra in relation to polynomials. If $f(x)$ is a polynomial, a *root* of the polynomial is a value of $x$ (either real or complex) such that $f(x) = 0$. When we consider a polynomial $f(x)$ with real coefficients and draw the graph of $y = f(x)$, the real roots of the polynomial appear as the $x$-coordinates of the points where the graph meets the $x$-axis. There is no such graphical meaning for complex roots. For instance, if $f(x) = 1 + x^2$ the roots are $\pm i$, where $i = \sqrt{-1}$. The graph of $y = 1 + x^2$ never goes below the line $y = 1$, and does not meet the $x$-axis.

Roots are related to factors as follows: If $r$ is a root, $x - r$ is a factor and vice versa. This means that we can write $f(x) = (x - r)F(x)$, where $F(x)$ is a polynomial of degree one less than that of $f(x)$. If $(x - r)^2$ is a factor but $(x - r)^3$ is not, the root $r$ is said to be a *double* root, or a root of *multiplicity* 2. Roots of higher multiplicity are defined similarly.

If a polynomial with real coefficients has a *nonreal* root, such roots occur in conjugate pairs, so that the total number of such roots is even. We emphasize that the truth of this statement depends on the fact that the coefficients of the polynomial are real numbers. For instance, if $2 + 3i$ is a root, so is $2 - 3i$. The product of the two factors corresponding to a pair of conjugate complex roots is a quadratic polynomial having the conjugate pair as roots. For example, if the conjugate roots are $2 \pm 3i$, the product of the factors is

$$[x - (2 + 3i)][x - (2 - 3i)] = (x - 2)^2 - (3i)^2 = (x - 2)^2 + 9$$
$$= x^2 - 4x + 13.$$

It is theoretically possible to factor a polynomial with real coefficients into factors of the type $(x - r)$ corresponding to a real root $r$ and quadratic factors arising as the product of factors corresponding to a pair of conjugate complex roots. Factors of either type may be repeated if there are roots of higher multiplicity. We say the factorization is "theoretically" possible; however, it may be practically difficult to find the roots.

The relation of roots to factors is made apparent by a long-division process. Let $q$ be any number; we divide $f(x)$ by $x - q$ until we get a constant remainder $R$. This is indicated in the form

$$\frac{f(x)}{x - q} = F(x) + \frac{R}{x - q},$$

where $F(x)$ is the quotient. We can write this in the form

$$f(x) = (x - q)F(x) + R. \tag{1}$$

**Example 2:**

$$\frac{3x^3 - 4x^2 + 2x + 5}{x - 1} = 3x^2 - x + 1 + \frac{6}{x - 1},$$

or $$3x^3 - 4x^2 + 2x + 5 = (x - 1)(3x^2 - x + 1) + 6.$$

Here $F(x) = 3x^2 - x + 1$ and $R = 6$. We leave the details of the long division to the student.

From (1) we see that $f(q) = 0 \cdot F(q) + R = R$. *Hence the constant remainder $R$ is the value of $f(x)$ when $x = q$.* In particular, then, $f(q) = 0$ if and only if $R = 0$, and in that case (1) shows that $x - q$ is a factor of the polynomial $f(x)$.

If we do the long division by the synthetic division process, this is frequently a good method of calculating the value $f(q)$; the arithmetic this way may be simpler than that involved in actual substitution of the value $x = q$.

**Example 3:** If $f(x) = 3x^4 - 5x^3 - 4x^2 - 2$, calculate $f(-2)$. The procedure is carried out as follows, with a zero put in for the coefficient of the first-degree term in $f(x)$:

$$\begin{array}{rrrrr|r} 3 & -5 & -4 & 0 & -2 & -2 \\ & -6 & 22 & -36 & 72 & \\ \hline 3 & -11 & 18 & -36 & 70 & \end{array}$$

The final entry in the last line is the desired value; in this case $f(-2) = 70$. The earlier entries are the coefficients of the quotient $F(x)$:

$$F(x) = 3x^3 - 11x^2 + 18x - 36.$$

In graphing a polynomial $f(x)$ with real coefficients we first look for the roots of the derivative $f'(x)$; these are the critical values of $f(x)$. As we pass across a real root of $f'(x)$, the derivative changes sign (from plus to minus or vice versa) if the multiplicity of the root is odd. But it *does not* change sign if the multiplicity is even. For example, $(x - 1)^2(x + 2)(x^2 + x + 1)$ changes sign as we cross $x = -2$, but not as we cross $x = 1$.

**Example 4:** Graph the equation

$$y = x^4 - 4x^3 - 2x^2 + 12x - 11$$

and locate the real roots of the polynomial.

The derivative is

$$\frac{dy}{dx} = 4x^3 - 12x^2 - 4x + 12$$
$$= 4(x^3 - 3x^2 - x + 3).$$

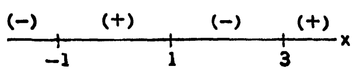
We see by inspection that $x = 3$ is a root of the derivative, so $x - 3$ is a factor. To get the quotient we use synthetic division.

$$\begin{array}{rrrr|r} 1 & -3 & -1 & 3 & 3 \\ & 3 & 0 & -3 & \\ \hline 1 & 0 & -1 & 0 & \end{array}$$

Thus $x^3 - 3x^2 - x + 3 = (x - 3)(x^2 - 1)$, and

$$\frac{dy}{dx} = 4(x - 3)(x - 1)(x + 1).$$

We diagram the sign of the derivative and tabulate the critical points of the




Sign of $\frac{dy}{dx}$

**Fig. 1-33**

graph, which are $(-1, -20)$, $(1, -4)$, and $(3, -20)$. With this much information we get a fairly good idea of what the curve looks like. But to locate the real roots it is necessary to tabulate points for some values of $x$ less than $-1$ and for some values greater than 3. So we make a table of values for the integers from $-2$ to 4. For convenience we use different scales on the two axes. It now appears that there are just two real roots, one between $-2$ and $-1$ (much nearer $-2$), and one between 3 and 4 (much nearer 4).

| $x$ | $y$ |
|---|---|
| $-2$ | 5 |
| $-1$ | $-20$ |
| 0 | $-11$ |
| 1 | $-4$ |
| 2 | $-11$ |
| 3 | $-20$ |
| 4 | 5 |

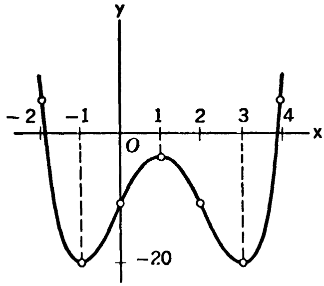



Fig. 1-34

## EXERCISES

1. Graph each polynomial by finding its critical points and using information provided by the sign of the derivative. Make a graph adequate to locate the real roots either at integers or between consecutive integers.
   (a) $y = x^3 - 3x^2 + 3$.
   (b) $y = x^3 - 3x$.
   (c) $y = 27x - x^3$.
   (d) $y = 2x^3 - 9x^2 + 12x - 3$.
   (e) $y = x^3 - 6x^2 + 12x - 5$.
   (f) $y = -x^3 + 6x^2 - 9x - 1$.
   (g) $y = \frac{1}{5}x^5 - \frac{4}{3}x^3$.
   (h) $y = x^4 + 4x^3$.
   (i) $y = \frac{1}{6}x^4 - x^2 + \frac{4}{3}x + \frac{5}{3}$.
   (j) $y = x^4 - 4x^3 + 16$.
   (k) $y = -x^4 + 2x^3 + 2x^2 - 1$.
   (l) $y = -3x^4 + 20x^3 - 42x^2 + 36x - 10$.
   (m) $y = \frac{3}{4}x^4 - 2x^3 + \frac{3}{2}x^2 - 3$.
   (n) $y = \frac{1}{6}x^6 + \frac{1}{5}x^5 - \frac{3}{2}x^4 - \frac{4}{3}x^3 + 4x^2 - 4$.

2. In each of the following cases $s$ is a polynomial function of $t$, defining a motion of a point on the $s$-axis. By graphing, with $t$ and $s$ in place of the usual $x$ and $y$, show when $s$ is increasing and when decreasing as $t$ increases, and hence describe how the point moves as $t$ increases from large negative values to large positive values.
   (a) $s = 96t - \frac{1}{2}t^3$.
   (b) $s = t^3 - 9t^2 + 15t - 7$.
   (c) $s = 6t^2 - 2t^3$.
   (d) $s = 64t^2 - 16t^4$.
   (e) $s = 8t^3 - 48t^2 + 72t$.

**3.** A right circular cylinder, radius of base $x$, is inscribed in a right circular cone whose base radius is 6 inches and whose height is 15 inches. Let $V$ be the volume of the cylinder, in cubic inches. Express $V$ as a polynomial in $x$ of third degree, and graph the function. Describe in words how $V$ varies as $x$ increases from 0 to 6.

**4.** From each corner of a square sheet of cardboard 24 inches on a side is cut a square of side $x$ inches. The edges of the sheet are then turned up to make a box. Express the volume $V$ of the box as a function of $x$ and graph the function.

**5.** Solve the problem corresponding to Exercise 4 if the cardboard sheet, instead of being square, is a rectangle 16 inches by 24 inches.

**6.** A crew of $x$ men works unloading boxes of manufactured goods from a freight car. The crew can unload $y$ cars per day, where $y = \frac{x^2}{25}\left(3 - \frac{x}{6}\right)$. Draw the graph of $y$ as a function of $x$ and discuss the effect of increasing the size of the crew on the number of cars unloaded per day. The formula is assumed to represent the facts of the situation for positive values of $x$ not exceeding 18.

**7.** The polynomial $x^4 - 4x^3 - 2x^2 + 12x - 11$ has a real root near $x = 4$, as we saw in Example 4. A better estimate of this root may be made as follows: Find the equation of the line tangent to the curve at $x = 4$ and locate the point where this tangent intersects the $x$-axis. This point is quite close to the desired root, but slightly to the right of it.

**8.** A sphere of radius $R$ is being filled with water. Let $x = h/R$, where $h$ is the depth of the water. Find a third-degree polynomial of which $x$ is a root when the sphere is $\frac{1}{4}$ filled (use the formula for the volume of a spherical segment). Draw the graph of this polynomial and estimate the value of the root in question. Observe that the tangent to the curve at $(1, -1)$ intersects the $x$-axis fairly near the desired root.

**9.** Solve Exercise 8 for the case in which the sphere is $\frac{3}{4}$ filled.

# CHAPTER II

# THE INVERSE OF DIFFERENTIATION

## 2-1 Some Fundamental Theory

In this section we consider a group of closely related theoretical matters pertaining to a function and its derivative. The main object of the section is to organize in logical order the sequence of ideas leading up to the theorem which is known as the *law of the mean,* and to explain two of the important uses of this theorem.

The contents of the section are drawn up in five items. The first three items are theorems, and the last two are corollaries of the third item. We go through all five items before discussing proofs of the theorems.

I. The first item is the following theorem.

THEOREM 2-A. *If $a$ and $b$ are two points on the x-axis, with $a < b$, and if a function $f$ is defined and continuous for each value of $x$ such that $a \leq x \leq b$, then, considering all the values $f(x)$ corresponding to these values of $x$, there is some $x$ for which the value of the function is algebraically largest and some $x$ for which the value of the function is algebraically smallest.*

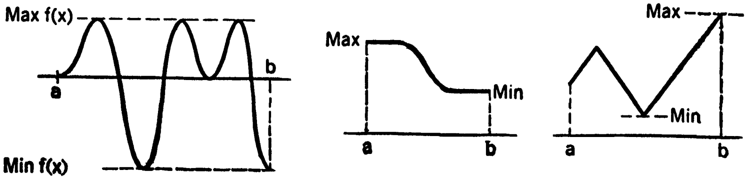



Fig. 2-1

The maximum of $f(x)$ may occur for more than one value of $x$. The important thing being asserted here is that there is *at least* one $x$ for which $f(x)$ is the maximum. Likewise for the minimum. Various possibilities for the occurrences of the maximum and minimum values are illustrated in Fig. 2-1.

II. The second item is also a theorem:

THEOREM 2-B. *If, among all the values attained by a function $f(x)$ as $x$ varies over a certain interval, an algebraic maximum is attained at a point $x_0$ inside the interval ("inside" means not at either end), and if the function is differentiable at this point $x_0$, then $f'(x_0) = 0$. The same is true for the case of an algebraic minimum.*

The case of a maximum of one of the types shown in Fig. 2-2 is ruled out in Theorem 2-B, because the function is not differentiable at $x_0$ in these cases.

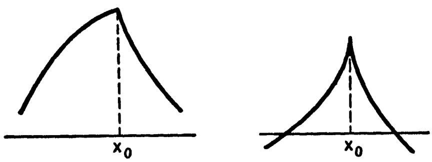


**Fig. 2-2**

III. Now suppose that we are given for consideration a function $f$ which is continuous at each point of an interval, including the end points, say for all $x$ such that $a \leq x \leq b$. Suppose also that the function is differentiable at each point *inside* the interval, so that the situations shown in Fig. 2-2 cannot occur. Construct the graph of $y = f(x)$ and draw the straight line joining the points $A$, $B$ of the graph corresponding to $x = a$, $x = b$ (see Fig. 2-3). The slope of this line is

$$\frac{f(b) - f(a)}{b - a}.$$

The *law of the mean* is the assertion that there is at least one value of $x$ between $a$ and $b$, say $x = X$, where $a < X < b$, such that

$$\frac{f(b) - f(a)}{b - a} = f'(X). \tag{1}$$

This means, geometrically, that the tangent to the graph when $x = X$ is parallel to the straight line through $A$ and $B$ (see Fig. 2-3). There may be more than one admissible value of $X$; the important thing is that there is *at least* one.

The name "law of the mean" comes from the use of "mean" in the sense of "average." The slope of the line $AB$ is a kind of average slope for the curve as $x$ goes from $a$ to $b$. If the curve does not coincide with the straight line $AB$, its slope will have various values, some more and some less than the slope of $AB$. But there will be at least one point of the curve where its slope is the same as the slope of $AB$.

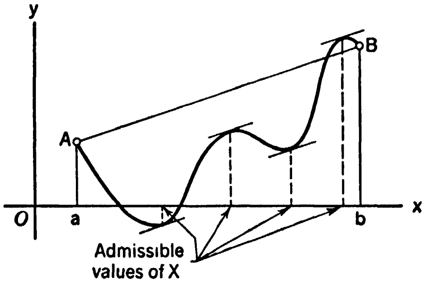



**Fig. 2-3**

Although the law of the mean has a clear geometrical interpretation, its uses do not depend so much on the geometrical meaning as upon the formula (1) which expresses the situation. This formula is often written in the form

$$f(b) - f(a) = (b - a)f'(X), \tag{2}$$

with the understanding that $X$ is *some* value of $x$ between $a$ and $b$. This formula may be applied to any function, with any choices of $a$ and $b$, provided the proper conditions are satisfied. For purposes of reference we state the law of the mean formally as a numbered theorem. In this statement we use $x_1$ and $x_2$ in place of $a$ and $b$.

THEOREM 2-C (*The law of the mean*). *If a function $f$ is continuous for each $x$ such that $x_1 \leq x \leq x_2$, where $x_1 < x_2$, and if it is differentiable for each $x$ such that $x_1 < x < x_2$, then there is at least one number $X$ such that $x_1 < X < x_2$ and*

$$f(x_2) - f(x_1) = (x_2 - x_1)f'(X). \tag{3}$$

We go on now to two items which are applications of the law of the mean.

IV. *If a function has a derivative which is positive at each point of an interval, the function is increasing on the interval.*

The truth of this statement can be seen at once with the help of the law of the mean. The conditions of the law of the mean are certainly fulfilled in this situation, because a differentiable function is continuous. Suppose $f'(x) > 0$ on an interval, and let $x_1$, $x_2$ be points of the interval with $x_1 < x_2$. Then (3) shows at once that $f(x_2) - f(x_1) > 0$, or $f(x_1) < f(x_2)$. This shows that $f$ is increasing on the interval. The student may contrast this argument with the one given in § 1-10. The present argument is shorter and simpler.

By a similar argument we may show that $f(x)$ decreases when $x$ increases if $f'(x)$ remains negative.

V. *If a function has a derivative which is zero at each point of an interval, the function is constant on that interval. That is, it has the same value for every $x$ on the interval.*

This also is an immediate corollary of the law of the mean. Suppose $f'(x) = 0$ for each $x$ on an interval, and suppose $x_1 < x_2$, where $x_1$ and $x_2$ are on the interval. Then the formula (3) shows that $f(x_1) = f(x_2)$. Hence $f(x)$ has the same value at all points of the interval.

In the remainder of this section we discuss Theorems 2-A, 2-B, and 2-C further, and consider the proofs.

In Theorem 2-A it is essential that the end points $x = a$, $x = b$ be included in the statement about where the function is continuous. If a function is continuous on an interval which does not contain both end points, there may not be any maximum or minimum value of the function on the interval. As $x$ approaches the end of the interval the function may pass through larger and larger values without ever reaching a maximum. This is illustrated by the case of $f(x) = 1/x$ as $x$ approaches 0 on the interval $0 < x \leq 1$. The function may also behave in more complicated ways.

There would be no difficulty in proving Theorem 2-A if we knew that the interval could be divided up into a finite number of parts (each part an interval) and that on each part the function either increased steadily or decreased steadily. For then we could examine the function values at the right end points of the intervals on which $f(x)$ increases and at the left end points of the intervals on which $f(x)$ decreases. The largest of the finite number of values so considered would be the maximum for the whole interval. This argument may be visualized by referring to the functions illustrated in Fig. 2-1. A similar argument can be made for the case of the minimum value. But this argument is not completely general, for it is a fact that continuous functions can be so complicated that this kind of analysis will not apply. However, the study of such complicated functions is beyond our present purposes. The usual place for studying a proof of Theorem 2-A is in a course of advanced calculus or other more advanced analysis, where the subject of continuity and the nature of the real number system are considered in detail. In this book we accept and use Theorem 2-A without further discussion.

There is no difficulty in proving Theorem 2-B. The important thing here is that the point $x_0$ of maximum or minimum value is inside the interval, and that the derivative at $x_0$ exists. The value of the derivative must then be 0. For suppose it were not 0. If the value were positive, the fact that

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) > 0$$

would mean that when $x$ is quite close to $x_0$ we have $f(x_0) < f(x)$ if $x_0 < x$ and $f(x) < f(x_0)$ if $x < x_0$. But then the value $f(x_0)$ could not be either a maximum or a minimum, as was supposed. The argument if $f'(x_0) < 0$ is essentially the same.

Our last consideration in this section is the proof of Theorem 2-C. For this we use both Theorem 2-A and Theorem 2-B. Instead of considering the function $f$ directly, we consider the amount by which the graph of $y = f(x)$ differs from the straight line $AB$ in Fig. 2-3. If the graph of $y = f(x)$ coincides with the line $AB$ the formula (1) is obviously true, so we have nothing to prove in that case. We let $g(x)$ = the directed distance $QP$, where $P$ is on the graph of $y = f(x)$ and $Q$ is the intersection of the line $AB$ and the line through $P$ parallel to the $y$-axis (see Fig. 2-4). The

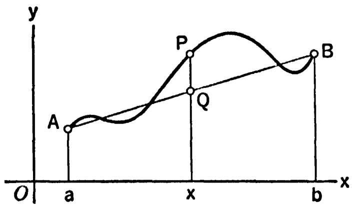



**Fig. 2-4**

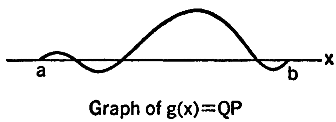



**Fig. 2-5**

graph of $y = g(x)$ will appear as in Fig. 2-5, the main feature being that $g(a) = g(b) = 0$. The actual formula for $g(x)$ is easily found. The equation of the line $AB$ in Fig. 2-4 is

$$y - f(a) = \frac{f(b) - f(a)}{b - a}(x - a),$$

and the $y$ here refers to $Q$. Thus $g(x) = f(x) - y$, or

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) - f(a). \tag{4}$$

Now $g(x)$ must attain both a maximum and a minimum value on the interval $a \le x \le b$, and since $g(a) = g(b) = 0$, either the maximum or the minimum and perhaps both will occur at a point inside the interval. Let $X$ be such a point. Then $g'(X) = 0$ (Theorem 2-B applied to $g$). But we see from (4) that

$$g'(X) = f'(X) - \frac{f(b) - f(a)}{b - a},$$

so $g'(X) = 0$ is the same as equation (1) or (2). In other words, the tangent to the graph of $y = g(x)$ is horizontal at the same time that the tangent to the graph of $y = f(x)$ is parallel to the line $AB$. This finishes the discussion of Theorem 2-C.

### EXERCISES

1. In each of the following situations there is no algebraically largest value among the considered values of the function. Explain which of the hypotheses of Theorem 2-A are not fulfilled. In each case we always have $f(x) < 1$, but values of $f(x)$ can be found as near 1 as one pleases.
   (a) The values of $f(x) = x^2$ corresponding to $0 \leq x < 1$.
   (b) The values, corresponding to $0 \leq x \leq 1$, of $f(x) = x - g(x)$, where $g(x) =$ the largest integer not exceeding $x$.
   (c) The values corresponding to $-1 \leq x \leq 1$ of $f(x)$, where by definition $f(x) = 1 - |x|$ if $0 < |x| \leq 1$ and $f(0) = 0$.
   (d) The values corresponding to all $x \geq 0$ of $f(x) = x^2/(1 + x^2)$.

2. (a) The function defined by $f(x) = |x|$ attains its absolute minimum value at $x = 0$. Why is not $f'(0) = 0$?
   (b) If $f(x) = 1 - x^{2/3}$, show that the absolute maximum of $f(x)$ for $x$ such that $-1 \leq x \leq 1$ is attained at a certain interior point of the interval. Try to compute $f'(x)$ at this point. Why is Theorem 2-B not applicable in this case?

3. If $f$ is defined when $0 < x < 2$ and differentiable at $x = 1$, with $f'(1) = \frac{1}{10}$, describe the relation between $f(x)$ and $f(1)$ when $x$ is sufficiently close to 1. Why is it impossible that $f(x) \leq f(1)$ for all $x$ such that $0 < x < 2$? Is it possible that $f(1) \leq f(x)$ for all such $x$? What would be the situation if $f'(1)$ were $-\frac{1}{10}$?

4. Suppose $f$ is differentiable when $a < x < b$, with $f'(x) \geq 0$ for each $x$ and $f'(x_0) > 0$ for a certain $x_0$. If $a < x_1 < x_0 < x_2 < b$, show that $f(x_1) < f(x_2)$.

5. Suppose $f$ is differentiable, with $f'(x) = 2$, when $a < x < b$, and that $f$ is continuous when $a \leq x \leq b$. Show that $f(x) = f(a) + 2(x - a)$ when $a \leq x \leq b$. Use V.

6. See if you can write out for yourself the proof of Theorem 2-C, with no help from the text except the guide furnished by Fig. 2-4 and Fig. 2-5.

## 2-2 Antiderivatives

In certain kinds of problems it is required to find all functions whose derivative is some specified function. We shall encounter problems of this kind in § 2-3 in connection with rectilinear motion, and in § 2-6 in connection with computation of areas.

Suppose $f(x)$ and $\phi(x)$ are defined on the same interval of the $x$-axis (perhaps even on the whole axis), and that $f'(x) = \phi(x)$. Then $f(x)$ is called an *antiderivative* of $\phi(x)$.

**Example 1:**
(a) $\frac{1}{4}x^4$ is an antiderivative of $x^3$.
(b) $\frac{3}{5}x^5 - \frac{7}{3}x^3 + 4x + 10$ is an antiderivative of $3x^4 - 7x^2 + 4$.

Finding an antiderivative of a given function is called *antidifferentiation*, because it is the inverse of differentiation. Some functions do not have antiderivatives, for there are functions which cannot be obtained by differentiating other functions. Antiderivatives are not unique; that is, if a function has one antiderivative, it actually has many antiderivatives, in fact, infinitely many. For if $f(x)$ is an antiderivative of $\phi(x)$, so is $f(x) + C$, where $C$ is any constant. This is because the derivative of the constant is zero, and

$$\frac{d}{dx}[f(x) + C] = \frac{d}{dx}f(x) + \frac{d}{dx}C = f'(x).$$

It is an important fact that all the antiderivatives of a given function can be obtained from a single one of the antiderivatives by adding constants to it.

THEOREM 2-D. *If $f(x)$ is any particular antiderivative of $\phi(x)$, the expression $f(x) + C$, where $C$ is an arbitrary constant, represents all possible antiderivatives of $\phi(x)$. This means that, if $g(x)$ is any other particular antiderivative of $\phi(x)$, there is some particular value of $C$ such that $g(x) = f(x) + C$.*

*Proof.* If $f(x)$ and $g(x)$ are particular antiderivatives of $\phi(x)$, let $F(x) = g(x) - f(x)$. Then

$$\frac{F(x+h) - F(x)}{h} = \frac{g(x+h) - g(x)}{h} - \frac{f(x+h) - f(x)}{h},$$

and so, if we take the limits as $h \to 0$,

$$F'(x) = g'(x) - f'(x) = \phi(x) - \phi(x) = 0.$$

But, as we saw in § 2-1, a function is constant on an interval if its derivative is always zero there. This was shown in item V, as a corollary of the law of the mean. Hence $F(x) = C$, where $C$ is some constant. But $g(x) = f(x) + F(x)$, so $g(x) = f(x) + C$, as asserted.

If $f(x)$ is a particular antiderivative of $\phi(x)$, the expression $f(x) + C$, with $C$ an arbitrary constant, is often called *the general antiderivative* of $\phi(x)$.

**Example 2:** The general antiderivative of $3x^2 - 5x - 10$ is $x^3 - \frac{5}{2}x^2 - 10x + C$.

Frequently we use a different terminology, and speak about *solving a differential equation*, instead of talking about antidifferentiation.

**Example 3:** The equation

$$\frac{dy}{dx} = 6x^3 - 2x + 1 \tag{1}$$

is called a *differential equation.* (It is a very special, simple kind of differential equation. Other types of differential equations are considered near the end of this book.) To solve the differential equation (1) means to find $y$ as a function of $x$ such that $dy/dx$ is given by (1). In other words, we must express $y$ as an antiderivative of $6x^3 - 2x + 1$. Finding the *general solution* of the differential equation is the same as finding the general antiderivative. The general solution of (1) is evidently

$$y = \frac{6x^4}{4} - \frac{2x^2}{2} + x + C = \frac{3}{2}x^4 - x^2 + x + C.$$

We know that differentiation of a polynomial applies to each term according to the rule

$$\frac{d}{dx}x^n = nx^{n-1}.$$

This procedure is reversed in antidifferentiation. *Hence, to get the general antiderivative of a polynomial, we raise the exponent of $x$ by 1 in each term, and divide the term by the new exponent. We then add an arbitrary constant to the whole expression.*

**Example 4:** The general antiderivative of $x^3 - 5x^2 + 7x - 2$ is

$$\frac{x^4}{4} - \frac{5x^3}{3} + \frac{7x^2}{2} - 2x + C.$$

Before learning how to find antiderivatives of functions other than polynomials we shall have to learn how to differentiate some other types of functions. Every time we learn a new differentiation formula we also get a new antidifferentiation formula, merely by looking at the differentiation formula in reverse.

In many problems we first find a general antiderivative, and then assign some particular value to the arbitrary constant; the assignment is made so as to satisfy some prescribed condition in the problem.

**Example 5:** Find the equation of a curve in the $xy$-plane, given that the slope of the curve satisfies the equation

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

and that the curve goes through the point $(1, -1)$.

Here we must begin by solving the differential equation (2). The general solution is

$$y = -\frac{x^2}{2} + C. \tag{3}$$

For each value of $C$ we get a curve whose slope satisfies equation (2). But only one of the curves goes through the point (1, −1). To find out which curve goes through the point, substitute $x = 1$, $y = -1$ in (3). The result is

$$-1 = -\frac{1}{2} + C, \quad \text{or} \quad C = -\frac{1}{2}.$$

Hence the equation of the required curve is

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

Several of the curves, including the one through (1, −1), are shown in Fig. 2-6.

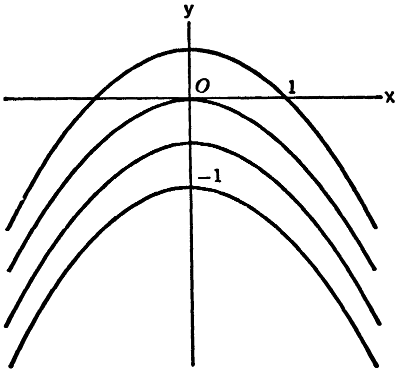



**Fig. 2-6**

## EXERCISES

**1.** Find the general antiderivative of each function. Check your answers.

(a) $x^2 - 4x + 13$.
(b) $3x^4 - 5x^3 - 4x^2 - 2$.
(c) $4x^3 - 12x^2 - 4x + 12$.
(d) $3x(x - 1)^2$.
(e) $\frac{2}{3}(x + 2)(x^2 - 1)$.
(f) $x^2(x^2 - 4)$.

**2.** Find the equation of the curve $y = f(x)$ through the given point, with slope at a typical point $(x, y)$ as given.

(a) Point (2, −3), slope $2x - 3$.
(b) Point (3, 10), slope $x^2$.
(c) Point (0, 0), slope $2x - 6x^3$.
(d) Point (2, 1), slope $(3x - 1)(x - 2)$.

**3.** Find $y$ as a function of $x$ from the given data.

(a) $\dfrac{dy}{dx} = x^2 - x$, $y = 1$ when $x = 0$.

(b) $\dfrac{dy}{dx} = -4x^3$, $y = 1$ when $x = -1$.

(c) $\dfrac{dy}{dx} = x(x - 1)(x - 2)$, $y = 4$ when $x = 2$.

**4.** The slope of a curve at $(x, y)$ is $Ax(x^2 - 1)$, where $A$ is some constant. The curve crosses the $x$-axis at $x = 3$ and it crosses the $y$-axis at $y = 2$. Find the equation of the curve.

5. Let $m$ denote the slope of a curve as a function of $x$. Find the equation of the curve from the data given.

(a) $\frac{dm}{dx} = -3$ for all $x$; $m = 4$ and $y = -1$ when $x = -1$.

(b) $\frac{dm}{dx} = 2x$; $m = 3$ and $y = 0$ when $x = 3$.

(c) $\frac{dm}{dx} = x - x^2$; $m = 2$ and $y = 6$ when $x = 2$.

(d) $\frac{dm}{dx} = -20x^3$; $y = 6$ when $x = 2$ and $y = -2$ when $x = -2$.

## 2-3 Rectilinear Motion

Rectilinear motion is motion in a straight line. By contrast, motion in a curved path is sometimes called curvilinear motion.

For mathematical purposes we often speak about a moving "particle." A particle is the ideal conception of an object which occupies but a single point of space. In discussing the motion of physical objects, such as spheres, cars, projectiles, and so on, we frequently ignore the size, shape, and orientation of the object and think of it as though it were a particle.

In this section we assume that all motions take place on the $s$-axis. The velocity of a particle whose coordinate is $s$ at time $t$ is

$$v = \frac{ds}{dt},$$

and its acceleration is

$$a = \frac{dv}{dt}.$$

Let us first consider a situation in which the acceleration is constant.

**Example 1:** A particle starts at $s = 10$ feet with velocity 5 feet per second, and moves with constant positive acceleration $a = 6$ feet per second per second. Find $v$ and $s$ as functions of $t$. In particular, find the values of $v$ and $s$ at 10 seconds after the start.

We take $t = 0$ at the start. We must distinguish carefully between the fact that $a = 6$ *at all times*, while $s = 10$ and $v = 5$ at the *particular instant* $t = 0$. The equation

$$a = \frac{dv}{dt} = 6$$

is a differential equation, which we solve by antidifferentiation:

$$v = 6t + A,$$

where $A$ is some constant. To find the value of $A$, put $t = 0$ and $v = 5$, as given. Then

$$5 = 6 \cdot 0 + A, \quad \text{or} \quad A = 5.$$

We now have

$$v = \frac{ds}{dt} = 6t + 5.$$

This also is a differential equation. Solving it, we get

$$s = \frac{6t^2}{2} + 5t + B = 3t^2 + 5t + B,$$

where $B$ is another constant. To find the value of $B$, put $s = 10$, $t = 0$. Then

$$10 = 3\cdot 0 + 5\cdot 0 + B, \quad \text{or} \quad B = 10.$$

We now have
$$s = 3t^2 + 5t + 10.$$

This formula describes the motion completely. When $t = 10$ we have

$$v = 65 \quad \text{and} \quad s = 360.$$

We used $A$ and $B$ for the constants obtained in antidifferentiation; these constants occur for the same reason that the arbitrary constant $C$ occurred in the examples in § 2-2.

Some students may be disposed to solve problems like that in Example 1 by using memorized formulas learned from physics, without resorting explicitly to the procedure of antidifferentiation. But our purpose here is to emphasize the value of the procedures of solving differential equations and evaluating the constants. The procedures afford us a single standard method for solving motion problems, and we are relieved of the necessity of memorizing an extensive list of formulas.

Bodies falling freely near the surface of the earth move with constant acceleration. (At least this is true as a satisfactory approximation for the study of most ordinary situations.) By *falling freely* we mean, strictly speaking, that the body falls in a vacuum. A body falling in air experiences some force due to air resistance, but this is negligible in many situations, and then we treat the situation as that of a freely falling body. The constant acceleration of a freely falling body is called the *acceleration due to gravity*. The numerical value of this constant is denoted by $g$. Its value depends on the units employed for distance and time; its values in the commonest systems of units are shown in the following table:

| Name of system* | Units of distance | Time | Approximate value of $g$ |
|---|---|---|---|
| British | feet | seconds | 32 |
| CGS | centimeters | seconds | 980 |
| MKS | meters | seconds | 9.8 |

* CGS stands for centimeter-gram-second; MKS stands for meter-kilogram-second.

The sign of the acceleration due to gravity is positive if the positive direction of the distance axis (in a vertical line) is downward; the acceleration is negative if the positive direction of the axis is upward.

**Example 2:** The $s$-axis extends positively upward with $s = 0$ at ground level. At time $t = 2$ seconds a boy leans out of a window 20 feet above the ground and tosses a ball straight upward with a speed of 24 feet per second. Find $s$ as a function of $t$ if $s$ is the number of feet the ball is above the ground at time $t$. Draw the graph.

The basic differential equation is

$$\frac{dv}{dt} = -32.$$

From it we find that

$$v = -32t + A,$$

where $A$ is some constant. Then

$$v = \frac{ds}{dt} = -32t + A,$$

and so

$$s = -16t^2 + At + B,$$

where $B$ is a second constant. The constants $A$ and $B$ are evaluated by using the information that $s = 20$ and $v = 24$ when $t = 2$. We have

$$24 = -64 + A,$$
$$20 = -64 + 2A + B.$$

Solving for $A$ and $B$ we find

$$A = 24 + 64 = 88, \quad B = 20 + 64 - 2(88) = -92.$$

Hence

$$s = -16t^2 + 88t - 92.$$

To draw the graph we use the derivative

$$v = \frac{ds}{dt} = -32t + 88,$$

which shows that $s$ is increasing if $t < \frac{11}{4}$ and decreasing if $t > \frac{11}{4}$. The value of $s$ when $v = 0$ is

$$s = -16\left(\frac{11}{4}\right)^2 + 88\left(\frac{11}{4}\right) - 92 = 29.$$

This is the highest point reached by the ball, 9 feet above where it was tossed by the boy. By tabulating the values of $s$ for $t = 1, 2, 3, 4$ we can draw a good graph (see Fig. 2-7). The curve is shown dotted before $t = 2$ and after the

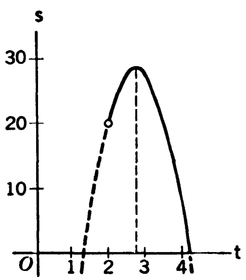



Fig. 2-7

| $t$ | $s$ |
|---|---|
| 1 | −20 |
| 2 | 20 |
| 3 | 28 |
| 4 | 4 |

value of $t$ for which the ball reaches the ground. This value can be found by setting $s = 0$ and solving a quadratic equation.

In some problems the acceleration may be constant, but unknown at the outset, the problem being such that the acceleration is determined by the information given. In this case the unknown constant acceleration is represented by a letter, and the work proceeds as though the literal constants were known until a point is reached where substitution of known data gives enough equations to solve for all the unknowns. Simultaneous equations may be involved.

**Example 3:** A car is braked to a stop with constant acceleration. The car stops 8 seconds after the brakes are applied, and travels 200 feet during this time. Find the acceleration, and the speed of the car when the brakes were applied.

For the solution let $k$ be the constant acceleration, and let $t = 0$ be the moment of applying the brakes. We have

$$\frac{dv}{dt} = k, \qquad v = kt + A,$$

where $A$ is some constant. Note that $v = A$ when $t = 0$, so $A$ is the velocity of the car at the time the brakes were applied. Then

$$\frac{ds}{dt} = kt + A, \qquad s = \frac{1}{2}kt^2 + At + B.$$

We take $s = 0$ at the point where the brakes were applied, and assume the car moves in the positive direction. Then $s = 200$ and $v = 0$ when $t = 8$. Substituting, we have

$$\begin{aligned} 0 &= \frac{1}{2}k\cdot 0 + A\cdot 0 + B = B && (t = 0,\ s = 0),\\ 200 &= 32k + 8A + B && (t = 8,\ s = 200),\\ 0 &= 8k + A && (t = 8,\ v = 0). \end{aligned}$$

Since $B = 0$, the last two equations become simultaneous equations for $A$ and $k$. Solving, we find

$$k = -\frac{25}{4} = -6\frac{1}{4} \quad \text{and} \quad A = 50.$$

Thus the constant acceleration is $-6\frac{1}{4}$ feet per second per second, and the initial velocity of the car was 50 feet per second.

The methods by which we solved the problems in Examples 1-3 apply just as well if the acceleration, instead of being constant, is any specified function of the time, provided that we can find an antiderivative of this function.

In converting from miles per hour to feet per second it is convenient to bear in mind that 60 miles per hour is the same as 88 feet per second.

## EXERCISES

1. A body is projected upward from the earth with velocity 196 meters per second. (a) Find $s$ as a function of $t$ in the MKS system, taking $t = 0$ and $s = 0$ at the start of the motion, with $s$ positive upward. (b) For how long, and how high, does the body rise?

2. (a) If air resistance were negligible, how long would it take an object, dropped out of an airplane at 20,000 feet above the sea, to strike the water? (b) Again neglecting air resistance, how long would it take a bullet to reach the ocean from the airplane, if it were fired downward with a muzzle velocity of 400 feet per second?

3. A ball is thrown vertically upward with an initial velocity of 30 meters per second from the edge of the top of a building 50 meters high. (a) Find the distance $s$ from the foot of the building up to the ball $t$ seconds later. (b) How long does it take the ball to hit the ground, assuming it misses the building on the way down?

4. A trailer is being pulled along a level street at the rate of 15 feet per second when it breaks loose from the car pulling it. If frictional forces produce a constant negative acceleration and the trailer stops after going 150 feet, what was the acceleration?

5. A block of ice slides down a chute with an acceleration of 18 feet per second per second. (a) If it is going 6 feet per second at a certain instant, how far does it slide in the next 3 seconds? (b) How far from the first position has it gone by the time its speed is 72 feet per second?

6. A boy at the top of a cliff throws a rock straight down and it hits the ground 189 feet below $2\frac{1}{4}$ seconds later. How fast did the boy throw the rock?

7. With what speed (in meters per second) must an arrow be shot up in order to fall back to its starting point in 10 seconds? How high will it rise?

8. A box slides down an inclined chute 140 feet long in $3\frac{1}{2}$ seconds, and arrives at the bottom with a velocity of 75 feet per second. Find the acceleration, and the velocity of the box at the top of the chute.

9. A ball is rolling down an incline with an acceleration of 24 feet per second per second. Find the equation of its motion, assuming that $s$ is measured down the incline, if $s = 11$ when $t = 2$ and $s = 120$ when $t = 3$.

10. A car, traveling 50 feet per second, starts to slow down with constant (negative) acceleration. In 3 seconds the speed is reduced to 20 feet per second. Find (a) the acceleration; (b) the time required to stop; and (c) the distance traveled while slowing to a stop.

11. An object is dropped out of a window 100 feet above the ground. At the same time another object is thrown straight down from a window 200 feet above the ground. Both objects reach the ground at the same time. Neglecting air resistance, find the initial velocity of the second object.

**12.** A car moving with constant acceleration has speed 15 miles per hour at $t = 0$, $s = 0$ and speed 60 miles per hour at $t = T$, $s = 660$. Find $T$ (in seconds), and the acceleration in feet per second per second.

**13.** (a) How much time does a train traveling 60 miles per hour require to stop with a constant negative acceleration of 8 feet per second per second? (b) How far does it travel in this time?

**14.** A particle moves with constant acceleration $a = k$ on the $s$-axis. Show that $v$ and $s$ are so related that $v^2 - 2ks$ is constant during the motion. In particular, if $v$ and $s$ are both 0 at some instant, then $v^2 = 2ks$. Suggestion: Find $v$ and $s$ in terms of $t$ and then eliminate $t$.

**15.** At the surface of the moon the acceleration of gravity is approximately $\frac{1}{6}$ that at the surface of the earth. At the surface of the sun the acceleration of gravity is approximately 28 times as great as at the surface of the earth. Suppose an object is thrown up from the earth's surface with enough velocity to carry it to a height of 7 feet (approximately the world's record for the high jump). How far up would it go (a) on the moon? (b) on the sun?

**16.** (a) What constant negative acceleration is required to bring a train to rest in $D$ feet if it is initially going $V$ feet per second? (b) Evaluate for $V = 132$, $D = 300$.

**17.** A ball is thrown straight up from ground level. A stop watch is started at a time whose relation to the instant the ball is thrown is not specified. When the watch registers $t = 2$ (seconds) the ball is 96 feet above the ground, and when the watch registers $t = 6$ the ball is 160 feet above ground. Answer the following questions after getting the equations which provide complete information about the motion. (a) What are the values of $t$ when the ball leaves the ground level and returns to it? (b) With what velocity was the ball thrown upward? (c) How high does the ball rise and for how many seconds does it rise?

**18.** Answer the same questions as in the preceding problem, the given data being that $s = 64$ when $t = 3$ and $v = -16$ when $t = 5$.

**19.** Assume that a man running the 100-yard dash maintains a constant acceleration for the first 32 yards and thereafter has 0 acceleration. What must the acceleration be if he is to run the race in 9.4 seconds?

**20.** A particle moves on the $s$-axis with acceleration $a = 2t - t^2$. Find $s$ as a function of $t$, assuming that $s = 0$ and $v = 0$ when $t = 0$. Draw the graph of $s$ as a function of $t$.

## 2-4 Parabolas

This section is devoted to study of the particular curves which are called parabolas.

*The Focus-Directrix Definition*

In a given plane consider a fixed straight line, denoted by $D$, and a fixed point not on the line, denoted by $F$. Consider the assemblage of all points $P$ in this plane which are such that the distance from $P$ to $F$ is the same as the perpendicular distance from $P$ to the line $D$. This assemblage of points forms a curve called *the parabola with directrix $D$ and focus $F$.* Figure 2-8 shows such a parabola, with a typical point $P$ on it. Evidently the parabola is symmetric with respect to the line through $F$ perpendicular to $D$. This line is called the *axis* of the parabola. The point $V$ where the parabola intersects its axis is called the *vertex* of the parabola; $V$ is midway between $F$ and $D$.

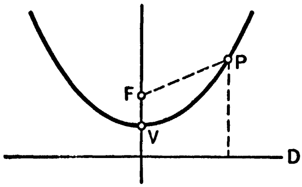



**Fig. 2-8**

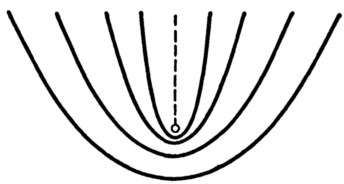


**Fig. 2-9**

If the directrix of the parabola is horizontal and the focus is above the directrix (as in Fig. 2-8), the parabola is said to open upward. If the focus were below the directrix the parabola would be said to open downward.

The effect on the appearance of the parabola of bringing the directrix closer to the focus is shown in Fig. 2-9, where all the parabolas have the same focus and all open upward along the same axis.

Our first objective will be to learn to identify parabolas by their equations when they are in the $xy$-plane and have their axes parallel either to the $x$-axis or the $y$-axis.

We consider first the case in which the directrix is parallel to the $x$-axis and the focus is above the directrix. One special case of this was considered in § 1-5 (Example 3). Suppose the vertex is the point $(a, b)$, and let $p$ be the distance between the focus and the directrix. Then the directrix is the line

$$y = b - \frac{p}{2},$$

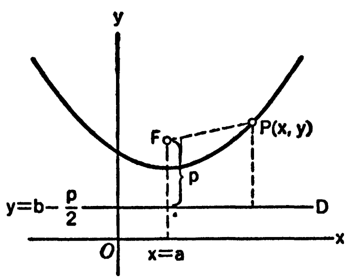



**Fig. 2-10**

and the coordinates of the focus are $x = a$, $y = b + p/2$ (see Fig. 2-10). The perpendicular distance from $P$ to $D$ is $y - (b - p/2)$, and the distance $PF$ is

$$\sqrt{(x-a)^2 + \left(y - b - \frac{p}{2}\right)^2}.$$

Hence, by the definition of the parabola, its equation is

$$\sqrt{(x-a)^2 + \left(y - b - \frac{p}{2}\right)^2} = y - b + \frac{p}{2}. \tag{1}$$

This equation can be greatly simplified in appearance. We square both sides of the equation, obtaining

$$(x-a)^2 + \left(y - b - \frac{p}{2}\right)^2 = \left(y - b + \frac{p}{2}\right)^2. \tag{2}$$

Next we write

$$\left(y - b - \frac{p}{2}\right)^2 = (y-b)^2 - p(y-b) + \frac{p^2}{4},$$

$$\left(y - b + \frac{p}{2}\right)^2 = (y-b)^2 + p(y-b) + \frac{p^2}{4}.$$

If these results are put into (2) and like terms are canceled from each side, we obtain the equation

$$(x-a)^2 - p(y-b) = p(y-b),$$

or

$$(x-a)^2 = 2p(y-b). \tag{3}$$

This is the simplified form of the equation (1). To prove that (3) is equivalent to (1) it is logically necessary to show that (1) can be derived from (3), i.e., that (1) holds if (3) does. In retracing steps from (3) to (1) we get as far as (2) simply by adding equal amounts to both sides of (3). In going from (2) to (1) we take square roots of each side of (2), and it is necessary to consider the other choice of sign, which would give us

$$-\sqrt{(x-a)^2 + \left(y - b - \frac{p}{2}\right)^2} = y - b + \frac{p}{2}. \tag{4}$$

This choice is incompatible with (3), however. For (4) implies $y - b + p/2 \leq 0$, or $y - b \leq -p/2$, so that $y - b$ would have to be negative. But (3) shows that $y - b$ cannot be negative. Hence we must reject (4). This completes the derivation of (1) from (3).

The student will find it worth while to memorize equation (3) as the equation of the parabola depicted in Fig. 2-10. The meaning of $(a, b)$ and $p$ must be kept in mind. In particular, if the vertex is at the origin, the equation has the very simple form

$$x^2 = 2py.$$

**Example 1:** Find the equation of the parabola with the line $y = -2$ as directrix and the point $(1, 5)$ as focus.

In this case the distance from focus to directrix is 7. The vertex is at $(1, \frac{3}{2})$. (The student should draw a figure and verify these assertions.) Hence the equation of the parabola is

$$(x - 1)^2 = 14\left(y - \frac{3}{2}\right).$$

### *The Latus Rectum*

If a line is drawn through the focus parallel to the directrix, the portion of this line cut off inside the parabola is called the *latus rectum* of the parabola. The student may observe from a figure that the length of the latus rectum is $2p$.

### *Other Standard Positions of Parabolas*

Suppose as before that the directrix is parallel to the $x$-axis, but this time let the focus be below the directrix, instead of above. The effect of this change on the equation of the parabola is easy to reckon. Using $(a, b)$ and $p$ just as before, we find that the equation of the parabola is

$$(x - a)^2 = -2p(y - b). \tag{5}$$

For purposes of memorization we note that the minus sign before the $2p$ in (5) is associated with the fact that the parabola opens downward (i.e., in the negative direction).

If the directrix of the parabola is parallel to the $y$-axis instead of to the $x$-axis, the equation of the parabola is obtained from (3) or (5) by interchanging $x$ and $y$ and also interchanging $a$ and $b$. The equation is

$$(y - b)^2 = 2p(x - a) \tag{6}$$

if the parabola opens to the right, and

$$(y - b)^2 = -2p(x - a) \tag{7}$$

if it opens to the left. In each case the vertex is at $(a, b)$ and $p$ is the distance between focus and directrix.

**Example 2:** $y^2 = -4x$ is the equation of the parabola with vertex at the origin, focus at $(-1, 0)$, and the line $x = 1$ for directrix. The student should draw a figure and mark the focus and directrix.

For many purposes the focus and directrix are not of primary interest, and frequently one deals with a parabola in terms of its vertex, its axis, and the direction in which it opens. If the vertex is at the origin and the axis of the parabola is along the $y$-axis, the equation of the curve has the form

$$x^2 = ky, \tag{8}$$

where $k$ is a constant (positive if the parabola opens upward, negative if it opens downward). If the axis of the parabola is along the $x$-axis and the vertex is at the origin, the equation has the form

$$y^2 = kx. \tag{9}$$

In either case the value of $k$ may be determined if we know a point on the parabola in addition to the vertex.

**Example 3:** A parabola with vertex at the origin is symmetric with respect to the $y$-axis and goes through the point $(3, -2)$. Find its equation.

We know the equation is of the form (8). To find $k$ we put $x = 3$, $y = -2$, getting $9 = -2k$, or $k = -\frac{9}{2}$. Thus the equation is

$$x^2 = -\frac{9}{2}y.$$

*Identification by Completing the Square*

Consider an equation of the form

$$y = Ax^2 + Bx + C, \tag{10}$$

where $A$, $B$, $C$ are constants, and $A \neq 0$. The equation (5) can be put in this form, and we shall see that this equation always represents a parabola with its axis parallel to the $y$-axis. We do this by putting (10) in the form (5). The procedure is to divide through by $A$ and then complete the square in the terms involving $x$:

$$\frac{1}{A}y = x^2 + \frac{B}{A}x + \left(\frac{B}{2A}\right)^2 + \frac{C}{A} - \frac{B^2}{4A^2}$$
$$= \left(x + \frac{B}{2A}\right)^2 + \frac{4AC - B^2}{4A^2}.$$

One should not memorize any formulas here; one need only recall the process of completing the square.

**Example 4:** Identify the parabola whose equation is $y = \frac{1}{2}x^2 + 3x + \frac{13}{2}$.

First we write

$$2y = x^2 + 6x \qquad +13,$$

leaving a space for the term which completes the square. The required term is 9, so we must compensate by subtracting it:

$$2y = x^2 + 6x + 9 + 13 - 9 = (x + 3)^2 + 4.$$

Finally we transpose and put the equation in the standard form (5):

$$(x + 3)^2 = 2y - 4 = 2(y - 2).$$

We now see that $a = -3$, $b = 2$, $p = 1$. The parabola has vertex at $(-3, 2)$, it opens upward, and the focus is one unit from the directrix.

A similar procedure applies to the equation

$$x = Ay^2 + By + C,$$

where $A \neq 0$. This is a general form for the equation of a parabola with axis parallel to the $x$-axis.

*Location of the Vertex by Calculus*

The vertex of the parabola with equation

$$y = Ax^2 + Bx + C$$

can be located also by using the formula

$$\frac{dy}{dx} = 2Ax + B$$

for the slope of the curve. The vertex is the point where the slope is 0. Hence it is at the point where $x = -B/2A$. The direction in which the parabola opens depends on the sign of $A$. If $A > 0$, the parabola opens upward, and if $A < 0$ it opens downward. To draw the graph quickly and easily, locate the vertex first of all. Then locate a pair of points on the curve symmetrically situated with respect to the axis of the parabola. With the vertex and these two points (at least if they are not too near the vertex) one may sketch the curve. Sometimes it is helpful or convenient to find where the curve crosses one or both of the coordinate axes.

**Example 5:** Locate the vertex of the parabola $y = 5 + 3x - 2x^2$ and sketch the curve.

Since

$$\frac{dy}{dx} = 3 - 4x = 0 \quad \text{if} \quad x = \frac{3}{4},$$

the vertex is at $x = \frac{3}{4}$. The corresponding value of $y$ is

$$y = 5 + 3\left(\frac{3}{4}\right) - 2\left(\frac{9}{16}\right) = \frac{49}{8}.$$

In this case the curve opens in the negative $y$-direction. Since $2x^2 - 3x - 5 = (2x - 5)(x + 1)$, the curve crosses the $x$-axis at $x = \frac{5}{2}$ and $x = -1$. For the graph see Fig. 2-11.

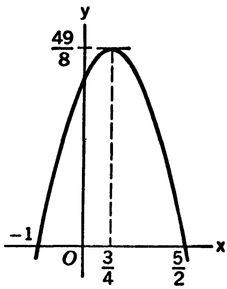


Fig. 2-11

One of the interesting facts about the parabola as a type of curve is

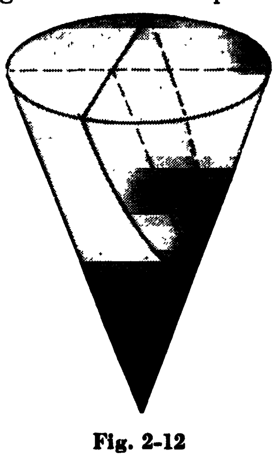


Fig. 2-12

that it is the curve obtained if the surface of a right circular cone is cut by a plane parallel to one of the elements of the cone (see Fig. 2-12). We shall not go into detail about this matter. There are also interesting uses of the parabola in optics and acoustics; these uses are based on certain facts about the tangents to a parabola. These matters are considered in the following section.

## EXERCISES

Draw each parabola when you work the exercise.

**1.** Find the equation of the parabola with:

(a) directrix $y = 0$ and focus $(0, 4)$;
(b) directrix $y = 3$ and focus $(0, 0)$;
(c) directrix $x = 0$ and focus $(-2, 0)$;
(d) directrix $x = -2$ and focus $(4, 0)$;
(e) vertex $(0, 0)$ and focus $(2, 0)$;
(f) vertex $(0, -3)$ and focus $(0, 0)$;
(g) vertex $(0, 0)$ and directrix $y = -2$;
(h) vertex $(4, 0)$ and focus $(0, 0)$.

**2.** Find the focus and directrix of each parabola.

(a) $y^2 = 8x$.
(b) $x^2 = y$.
(c) $y = -2x^2$.
(d) $4x^2 + 9y = 0$.
(e) $3y^2 = 4x$.
(f) $5x + 3y^2 = 0$.
(g) $x^2 = 4my$ $(m > 0)$.
(h) $y^2 = kx$ $(k < 0)$.
(i) $y^2 = 2px - p^2$ $(p > 0)$.
(j) $x^2 = 2py + p^2$ $(p > 0)$.

**3.** Find the vertex, focus and directrix of each parabola.

(a) $y^2 = 2x + 4$.
(b) $x^2 = 2x + 4y$.
(c) $x^2 + 2x = 4y - 9$.
(d) $y^2 - 6y + 24x = 87$.
(e) $x^2 + 8x - 16y = 16$.
(f) $y^2 - 2y + 10x = 44$.
(g) $x^2 + 8y + 4x - 20 = 0$.
(h) $x^2 - 12x + 12y + 48 = 0$.
(i) $x^2 + 8x + 10y = 34$.
(j) $y^2 + 2y + 29 = 7x$.
(k) $y^2 = 12x - 8y - 30$.
(l) $3x^2 - 6x - 4y + 11 = 0$.
(m) $4y^2 + 3x - 24y + 42 = 0$.
(n) $3y^2 = 6y - 5x - 13$.

**4.** Find the equation of the parabola with vertex at $(0, 0)$, axis along the $x$-axis, and the curve

(a) through $(6, -4)$;
(b) through $(-2, 4)$;
(c) cutting the line $3x + 4y = 18$ at $x = 2$;
(d) cutting the parabola $2y = x^2$ at $x = \frac{5}{2}$.

**5.** Find the equation of each parabola, from the indicated diagram.
(a) Fig. 2-13. (b) Fig. 2-14.
(c) Fig. 2-15. (d) Fig. 2-16.

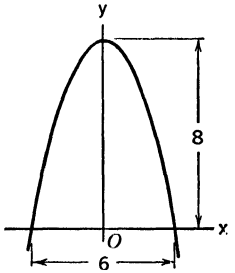



**Fig. 2-13**

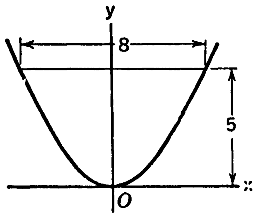



**Fig. 2-14**

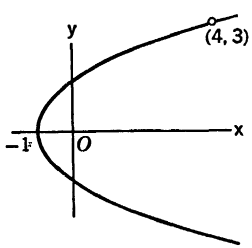



**Fig. 2-15**

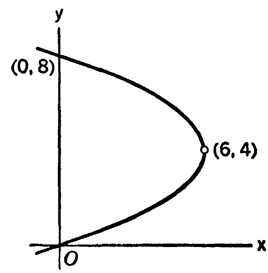



**Fig. 2-16**

**6.** A parabola has its axis on the $y$-axis, and it goes through the points (2, 3), (−1, −2). Find its equation.

**7.** A parabola with axis parallel to the $y$-axis goes through (0, 0), (1, 0), and (3, 6). Find its equation.

**8.** Find the equation of the parabola through the points (1, 0), (−7, 0), (−3, 2), given that its axis is parallel to one of the coordinate axes.

**9.** Find the equation of the parabola with axis parallel to the $y$-axis, if it goes through:
(a) (1, 2), (3, 4), and (6, 3);
(b) (−1, 1), (1, 0), and (8, 2);
(c) (2, 1), (3, 3), and (5, −1);
(d) (−2, 1), (1, −3), and (2, −2).

**10.** Find the equation of the parabola with axis parallel to the $x$-axis, if it goes through:
(a) (−1, 5), (7, 9), (2, 7);
(b) ($-\frac{7}{3}$, 5), (3, −3), (1, 3);
(c) (0, 3), (3, −6), (8, 9);
(d) (9, 0), (1, 8), (6, −2).

**11.** Figure 2-17 shows a parabolic arch 40 feet wide at the base, and 36 feet high. Find the vertical clearance under the arch at 5-foot intervals across the base.

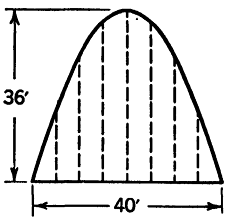



**Fig. 2-17**

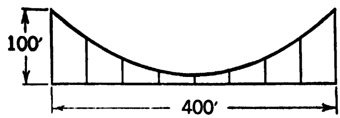



**Fig. 2-18**

**12.** Figure 2-18 shows a roadway 400 feet long held up by a parabolic cable. The cable is 100 feet above the roadway at the ends, and 4 feet above at the center. Find the lengths of the vertical supporting cables at 50-foot intervals along the roadway.

**13.** The surface of the roadway over a stone bridge follows a parabolic curve. The span of the bridge is 60 feet, and the road surface is 1 foot higher in the middle than at the ends. How much higher than the ends is a point of the roadway 15 feet from an end?

**14.** A focal chord of a parabola is the segment cut by the parabola from a straight line through the focus. Show that a focal chord is twice as long as the distance from its mid-point to the directrix.

**15.** A ball is thrown so that it starts out horizontally. It drops 1 foot in going 10 feet horizontally. (a) How far does it drop in going 20 feet horizontally? (b) How far horizontally does it go in dropping 9 feet? Assume that the path is a parabola with axis vertical.

**16.** A stone is tossed from a point 6 feet above the ground. It rises 3 feet in going the first 4 feet horizontally, and rises another $\frac{3}{4}$ foot in going the next 2 feet horizontally. How far does it go horizontally before it hits the ground? Assume that the path is a parabola with axis vertical.

**17.** A certain street is 30 feet wide. In cross section the road surface is a parabola with axis vertical. One side of the street is 6 inches higher than the other, and the middle of the street is $7\frac{1}{2}$ inches higher than the lower side. Find the greatest height of the road surface, and the heights at intervals of 5 feet across the section (see Fig. 2-19 in which the vertical scale is exaggerated).

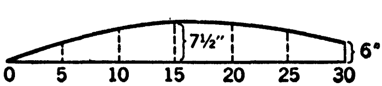



**Fig. 2-19**

**18.** A particle moves on the $s$-axis with constant acceleration $-g$ (a freely falling body with $s$-axis positive upward). When $t = 0$, $s = s_0$ and $v = v_0$. Obtain the equation expressing $s$ as a function of $t$. With $t$ and $s$ in place of $x$ and $y$ show that the graph is a parabola with axis parallel to the $s$-axis and vertex at $t = v_0/g$. How high is the vertex above the point $(0, s_0)$, assuming $v_0 > 0$?

## 2-5 Tangents to Parabolas

*The Optical Property of a Parabola*

Probably the most interesting thing about a parabola is the geometrical property which is responsible for the use of the parabola in lamps and in the reflectors of telescopes. The property can be stated as follows: *Let F be the focus of a parabola, let P be any point on the curve, and let T be the tangent to the parabola at P. Draw a line through P parallel to the axis of the parabola. Then this line and the focal radius FP make equal angles with T.* For the diagram in Fig. 2-20 this amounts to saying that $\alpha = \beta$; in

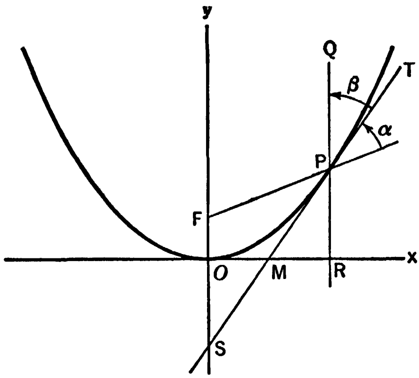



**Fig. 2-20**

other words, $T$ bisects the angle between the line $FP$ produced and the line $PQ$ parallel to the axis.

To prove this we must first of all represent the parabola by its equation, and find the slope of the curve at $P$. For the location shown in Fig. 2-20 the curve will have an equation $x^2 = 2py$, or

$$y = \frac{1}{2p}x^2, \tag{1}$$

where the focus $F$ is at $(0, p/2)$. If $P$ is the point $(x, y)$, the slope at $P$ is

$$\frac{dy}{dx} = \frac{1}{2p}2x = \frac{x}{p}.$$

This is the tangent of the angle of inclination of $T$, or the cotangent of the complementary angle $QPT$; that is

$$\operatorname{ctn} \beta = \frac{x}{p}, \quad \text{or} \quad \tan \beta = \frac{p}{x}. \tag{2}$$

The slope of $PF$ is

$$\frac{y - \frac{p}{2}}{x - 0} = \frac{\frac{x^2}{2p} - \frac{p}{2}}{x} = \frac{x^2 - p^2}{2px}.$$

Therefore, by the standard method for dealing with the angle between two lines, we have

$$\tan \alpha = \frac{\frac{x}{p} - \frac{x^2 - p^2}{2px}}{1 + \frac{x}{p}\frac{x^2 - p^2}{2px}} = \frac{2x^2 - x^2 + p^2}{2px} \cdot \frac{2p^2}{2p^2 + x^2 - p^2}.$$

On simplification this becomes

$$\tan \alpha = \frac{p}{x}. \tag{3}$$

Comparing (2) and (3), we see that $\alpha = \beta$, as asserted.

The optical interpretation of what we have just proved is this: Suppose the parabola represents a reflecting surface. If a ray of light $QP$ comes in parallel to the axis of the parabola and strikes the reflector at $P$, it will be reflected along the line from $P$ to the focus $F$. This is because of the property of the parabola and the physical law that the angle of incidence equals the angle of reflection. Light coming from a great distance (e.g., from the stars) is in parallel rays. Hence a parabolic reflector pointed toward a star will bring all the rays together at the focus. If the reflector is used with a source of light at $F$, the light will be reflected out in a bundle of parallel rays.

### *The Equation of a Tangent*

Now let us consider the parabola in Fig. 2-20, with equation (1), and suppose that $P$ is a typical point $(x_0, y_0)$ on the curve. The slope at $P$ is $x_0/p$, as we have already seen by differentiating (1). Hence the tangent at $P$ has the equation

$$y - y_0 = \frac{x_0}{p}(x - x_0). \tag{4}$$

For certain purposes it is convenient to rewrite (4) in a different form. We expand the right side and replace $x_0^2$ by $2py_0$, using (1):

$$y - y_0 = \frac{x_0 x}{p} - \frac{x_0^2}{p} = \frac{x_0 x}{p} - 2y_0.$$

We then multiply through by $p$, transpose terms, and collect. The result is

$$x_0x = p(y + y_0). \tag{5}$$

This equation may be used in solving various problems dealing with tangents to the parabola.

**Example:** Show that the tangent at an end of the latus rectum intersects the axis of the parabola on the directrix, and that it makes an angle of 45° with the axis.

The latus rectum is the chord of the parabola through $F$ parallel to the directrix. The ends of the latus rectum are the points $(\pm p, p/2)$. Taking $x_0 = p$, $y_0 = p/2$ in (5) we obtain

$$px = py + \frac{p^2}{2}, \quad \text{or} \quad y = x - \frac{p}{2}.$$

Thus the tangent has slope 1 and $y$-intercept $-p/2$. It follows at once that the tangent does what is asserted in the statement of the example.

## EXERCISES

**1.** Find the tangent to
(a) $x^2 = 4y$ at $(4, 4)$.
(b) $y = x^2$ at $(2, 4)$.
(c) $x^2 = 9y$ at $(3, 1)$.
(d) $x^2 = -4y$ at $(2, -1)$.
(e) $3y = -x^2$ at $(3, -3)$.
(f) $x^2 + 9y = 0$ at $(6, -4)$.
Draw each parabola and the tangent in question.

**2.** Find the equation which corresponds to (5) for the parabola $x^2 + 2py = 0$.

**3.** Find the tangent to:
(a) $x^2 = 9y$ with slope $\frac{2}{3}$.
(b) $x^2 = 4y$ with slope 10.
(c) $x^2 = 8y$ crossing the $x$-axis at $(4, 0)$.
(d) $x^2 = 6y$ crossing the $y$-axis at $(0, -6)$.
Draw each parabola and the tangent in question.

**4.** (a) Consider the tangent to the parabola $x^2 = 4y$ at the point $(4, 4)$. Let $A$ be the point where the tangent crosses the $x$-axis. Show that the line from the focus to $A$ is perpendicular to the tangent. (b) Show that this assertion is still true if we consider the tangent at any point (not the vertex) on the parabola $x^2 = 2py$.

**5.** (a) Consider the tangent to the parabola $x^2 = 2y$ at the point $(4, 8)$. Let $A$ and $B$ be the points where the tangent crosses the directrix and latus rectum produced, respectively. Show that $A$ and $B$ are equidistant from the focus. (b) Show that this assertion is still true if we consider the tangent at any point (not the vertex) on the parabola $x^2 = 2py$.

**6.** When referred to a suitable coordinate system, the engineer's drawing of the curved part of a roof truss is that part of the parabola $x^2 + 48y = 576$

above the $x$-axis. A brace is shown, running from one of the points on the parabola where $y = 6$, and at right angles to the tangent there, to the $x$-axis. Find the length of this brace.

7. In Fig. 2-20 show that $SO = RP$ and that $OM = MR$.

8. A line is drawn through a point $P$ on a parabola, perpendicular to the tangent there. The line cuts the axis of the parabola at $N$. Show that the focus is equidistant from $P$ and $N$.

9. A tangent to the parabola at $P$ (not the vertex) cuts the directrix at $A$. Show that angle $AFP$ is a right angle (where $F$ is the focus).

10. A straight line is drawn through the focus of a parabola, intersecting the curve at $(x_0, y_0)$ and $(x_1, y_1)$. Prove that the tangents to the parabola at these points intersect at right angles on the directrix. *Suggestion:* Use Fig. 2-20, taking $P$ to be $(x_0, y_0)$, and putting in the other point. For each tangent find the $x$-coordinate of the point where it crosses the directrix. Then use the fact that $(x_0, y_0)$, $F$, and $(x_1, y_1)$ are in a straight line to prove that the two values of $x$ so found are equal.

11. Two parabolas have the same focus and the same axis, but their vertices are on opposite sides of the focus (but not necessarily the same distance from it). Show that the curves cut each other at right angles. For convenience take the origin as focus and the $y$-axis as axis of symmetry.

## 2-6 The Definition of Area Under a Curve

The area concept requires careful definition. A square one unit on a side has, by definition, *one square unit* of area. A rectangle 3 units long and 2 units wide can be divided into 6 unit squares, and hence has 6 square units of area. For rectangles in general we must resort to the use of smaller squares as well as unit squares, and we are led to the formula for the area of a rectangle:

$$\text{area} = \text{length times breadth}.$$

Thus a rectangle of dimensions $3\frac{1}{2}$ by $2\frac{1}{4}$ can be divided up into 6 unit squares and 30 smaller squares $\frac{1}{4}$ unit on a side. See Fig. 2-21. Since 16 of these latter squares fill a unit square, the area of one then is $\frac{1}{16}$ of a square unit. The total number of square units of area in the rectangle is therefore

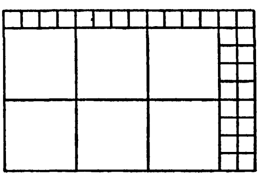


Fig. 2-21

$$6 + 30(\tfrac{1}{16}) = 7\tfrac{7}{8}.$$

This is the same as the result given by the formula:

$$\text{area} = (3\tfrac{1}{2})(2\tfrac{1}{4}) = 7\tfrac{7}{8}.$$

If the dimensions of the rectangle are not commensurable with the unit of length, i.e., if one or both dimensions is not expressible as a ratio of integers, the procedure is more complicated, and involves essentially a limiting process similar to that involved in the calculation of lengths incommensurable with a unit length. For instance, the length of the diagonal of a unit square has length $\sqrt{2}$; the number $\sqrt{2}$ is not expressible as a ratio of integers, and must be regarded as the limit of a succession of ratios of integers, for example:

$$\frac{14}{10}, \quad \frac{141}{100}, \quad \frac{1414}{1000}, \quad \frac{14{,}142}{10{,}000}, \quad \ldots.$$

In what follows we take the formula for area of a rectangle as fundamental, and we seek to define and compute the areas of other kinds of plane figures. We are especially interested in areas which are bounded, at least in part, by curved lines.

The basic principle is the following: we obtain estimates for the area of a figure by covering it as nearly exactly as we can by nonoverlapping rectangles. We then add up the area of these rectangles and use the sum as an approximation for the area of the given figure. The approximation will be too small if all the rectangles lie in the figure and if the figure is not entirely covered. The approximation will be too large if the rectangles cover the figure completely and something outside the figure in addition. *If we can get approximations, both too small and too large, which come as close as we please to the same number, this number is what we define to be the area of the given figure.*

In order to clarify the meaning of this definition, let us discuss the area for a figure partly bounded by a parabola. Let it be required to compute the area of the figure bounded by the arc of the parabola $y = \frac{1}{4}x^2$ from $x = 0$ to $x = 4$, the $x$-axis from $x = 0$ to $x = 4$, and the line $x = 4$ from

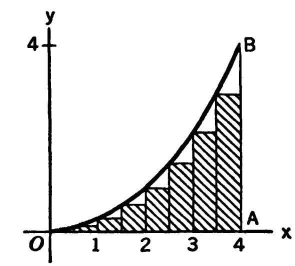



**Fig. 2-22**

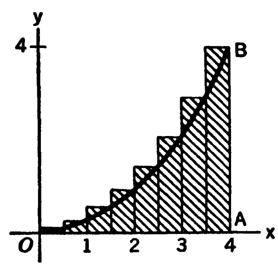



**Fig. 2-23**

$y = 0$ to $y = 4$. This is the triangle-like figure $OAB$ in Fig. 2-22 and Fig. 2-23. Figure 2-22 shows a set of rectangles which give a too small approximation to the area of the figure, while the set of rectangles in Fig. 2-23 give an approximation which is too large. These figures are obtained by dividing the base line $OA$ into 8 equal subintervals. To obtain better and better approximations, we divide $OA$ into $n$ equal parts and construct figures corresponding to those in Fig. 2-22 and Fig. 2-23. Then we see what happens to these approximations as $n$ gets larger and larger.

For $n$ equal parts of the base $OA$ the points of subdivision on the $x$-axis are at

$$0, \quad \frac{4}{n}, \quad \frac{8}{n}, \quad \frac{12}{n}, \quad \ldots, \quad \frac{4n-4}{n}, \quad 4.$$

The width of the base of each rectangle is $4/n$. In Fig. 2-22 the height of a typical rectangle is $y = \frac{1}{4}(4k/n)^2$, corresponding to $x = 4k/n$ at the lower left corner of the rectangle. Hence the sum of the areas of the rectangles in Fig. 2-22 is

$$A_n = \frac{1}{4}\left[\left(\frac{4}{n}\right)^2 + \left(\frac{8}{n}\right)^2 + \cdots + \left(\frac{4n-4}{n}\right)^2\right]\frac{4}{n}.$$

On simplification this becomes

$$A_n = \frac{16}{n^3}\left[1^2 + 2^2 + \cdots + (n-1)^2\right]. \tag{1}$$

For Fig. 2-23 the calculations are similar, and the sum of the areas of all the rectangles is

$$S_n = \frac{16}{n^3}\left[1^2 + 2^2 + \cdots + n^2\right]. \tag{2}$$

For the case $n = 8$ illustrated in the two figures we have

$$A_8 = \frac{1}{32}\left[1^2 + 2^2 + \cdots + 7^2\right] = \frac{35}{8},$$

$$S_8 = \frac{1}{32}\left[1^2 + 2^2 + \cdots + 8^2\right] = \frac{51}{8}.$$

In the general case we need a formula for computing sums of squares of consecutive integers. There is such a formula; it is

$$1^2 + 2^2 + \cdots + p^2 = \frac{p(p+1)(2p+1)}{6}, \tag{3}$$

where $p$ can be any positive integer. A discussion of this formula will be found in the paragraph on mathematical induction at the end of this section. If we put $p = n - 1$ in (3) we see from (1) that

$$A_n = \frac{16}{n^3} \cdot \frac{(n-1)n(2n-1)}{6} = \frac{8}{3} \cdot \frac{2n^2 - 3n + 1}{n^2}.$$

We can rewrite this as

$$A_n = \frac{8}{3}\left(2 - \frac{3}{n} + \frac{1}{n^2}\right),$$

and it is then evident that $A_n$ approaches $\frac{16}{3}$ as we make $n$ larger and larger. Likewise, putting $p = n$ in (3) and using (2), we see that

$$S_n = \frac{8}{3}\left(2 + \frac{3}{n} + \frac{1}{n^2}\right).$$

The difference between $S_n$ and $A_n$ is

$$S_n - A_n = \frac{16}{n};$$

this difference approaches 0 as $n$ increases. Hence $A_n$ and $S_n$ both approach the same limit, $\frac{16}{3}$. By our fundamental principle about the definition of areas we conclude that the area of the curved figure $OAB$ is $\frac{16}{3}$ square units. We observe, incidentally, that this is one third of the area of a square 4 units on a side.

Next we shall indicate how the foregoing procedure can be carried out, at least in theory, to obtain the area bounded by the lines $x = a$, $x = b$, the $x$-axis, and the graph of $y = f(x)$, where $f$ is a function which is continuous and never negative when $a \le x \le b$. Figure 2-24 and Fig. 2-25

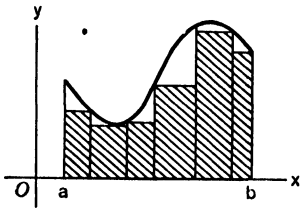



**Fig. 2-24**

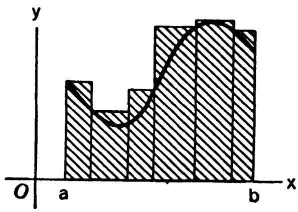



**Fig. 2-25**

show how we obtain minor (too small) and major (too large) approximations to the area in question by the construction of certain sets of rectangles. It is not essential that the rectangles all be of equal width. For the minor approximation the height of a rectangle is the minimum value of $f(x)$ for values of $x$ in the subinterval forming the base of the rectangle. For the major approximation we use for the height of the rectangle the maximum value of the function on the same subinterval.

When we increase the number of subdivision of the base line in such a way that the greatest of the widths of the rectangles approaches zero, it turns out that the minor and major approximations approach a common limiting value which is, by definition, the area between the curve $y = f(x)$ and the $x$-axis from $x = a$ to $x = b$. The detailed justification of this statement involves the use in a fundamental way of the fact that the function $f$ is assumed to be continuous. We omit these details, the elaboration of which is better postponed to a later stage in the student's progress.

The purpose of this section is purely to explain the concept of the area of a plane figure. The definition of area of a figure indicates a procedure for calculating the area approximately as the sum of the areas of a number of rectangles. To get the area exactly by direct application of this procedure, we have to carry out a limiting process as the number of rectangles increases. The details of this limiting process depend upon the particular curve we are considering. In general these details may be long and complex. Hence it is fortunate that the direct limiting procedure can in practice usually be avoided. It turns out that there is a connection between antidifferentiation and the finding of areas. This connection is explained in § 2-7. It is worthy of emphasis, however, that approximate answers are often useful in engineering and physics, and that fairly close approximations to the area under a curve can be made without a limiting process. One merely needs a carefully drawn graph on a fairly large scale. A suitable set of rectangles can then be drawn in and the sum of their areas reckoned by direct measurement of dimensions on the graph.

Instruments have been devised for calculating area by applying the instrument to the graph. The interested student may read about one such instrument, the planimeter, in an encyclopedia.

The average of a major and minor approximation may be a pretty accurate estimate of the exact area. For instance, the average of $A_8$ and $S_8$ in the case worked out earlier is $\frac{43}{8} = 5\frac{3}{8}$. The exact area is $\frac{16}{3} = 5\frac{1}{3}$, or $\frac{1}{24}$ less.

*Uses of Mathematical Induction*

Our need for formula (3) makes it appropriate for us to insert here some remarks about mathematical induction, for it is by mathematical induction that we may verify the fact that (3) is true for all positive integers $p$.

Mathematical induction is the name which is given to a principle of reasoning which is frequently used in proving general assertions involving positive integers in some manner. The basis for the principle is the fact that if we start with the number 1 and obtain other numbers by successive additions of 1, the numbers so obtained (by indefinite continuation of the process) comprise all positive integers:

$$\begin{aligned} 1 &= 1 \\ 1 + 1 &= 2 \\ 2 + 1 &= 3 \\ 3 + 1 &= 4 \\ \cdot\quad &\quad\cdot \\ \cdot\quad &\quad\cdot \\ \cdot\quad &\quad\cdot \end{aligned}$$

Now suppose that we have a sequence of propositions, one proposition associated with each positive integer. Let us symbolize these propositions by $P_1$, $P_2$, $P_3$, $\cdots$. The principle of mathematical induction then says: If $P_1$ is true, and if for every $k$, from the assumption that $P_k$ is true, we can deduce that $P_{k+1}$ is true, then $P_n$ is true for every positive integer $n$. For $P_1$ is true; this implies that $P_2$ is true, which implies that $P_3$ is true, and so on.

We illustrate by letting $P_k$ be the proposition expressed in the formula

$$1^2 + 2^2 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6}.$$

In particular $P_1$ is the proposition

$$1^2 = \frac{1 \cdot 2 \cdot 3}{6},$$

which is certainly true. Now does the truth of $P_k$ imply that of $P_{k+1}$? Let us see. The proposition $P_{k+1}$ is just like $P_k$ except that we must replace $k$ by $k + 1$. Since $2(k + 1) + 1 = 2k + 3$, $P_{k+1}$ is the proposition

$$1^2 + 2^2 + \cdots + k^2 + (k+1)^2 = \frac{(k+1)(k+2)(2k+3)}{6}.$$

We assume, tentatively, that $P_k$ is true. Then

$$\begin{aligned} 1^2 + 2^2 + \cdots + k^2 + (k+1)^2 &= \frac{k(k+1)(2k+1)}{6} + (k+1)^2 \\ &= \frac{(k+1)[k(2k+1) + 6(k+1)]}{6} \\ &= \frac{(k+1)(2k^2 + 7k + 6)}{6} \\ &= \frac{(k+1)(k+2)(2k+3)}{6}. \end{aligned}$$

Thus we deduce that $P_{k+1}$ is true if $P_k$ is. The principle of mathematical induction then implies that $P_n$ is true for all positive integers $n$.

Mathematical induction does not explain how the formula (3) was discovered in the first place. It may have been by ingenuity and conjecture, as is often the case with mathematical discoveries.

## EXERCISES

1. Find the area between the parabola $y = \frac{1}{2}x^2$ and the $x$-axis, from $x = 0$ to $x = 3$, by the method employed in the text.

2. Find the area between the $x$-axis and the line $by = hx$ ($b$ and $h > 0$), from $x = 0$ to $x = b$, by the method employed in the text. Use the formula

$$1 + 2 + \cdots + p = \frac{p(p+1)}{2}.$$

Partial answer: The minor approximation with $n$ equal subintervals is

$$A_n = \frac{hb}{n^2}[1 + 2 + \cdots + (n - 1)].$$

3. Verify by mathematical induction the truth of the formula for $1 + 2 + \cdots + p$ in the foregoing exercise.

4. Plot the curve $y = x^3 - 6x^2 + 9x + 1$ for $0 \leq x \leq 4$. Calculate the minor and major approximations to the area between the curve and the $x$-axis, using the method illustrated in Fig. 2-24 and Fig. 2-25, with 8 equal subintervals of the base. In this case it happens that the average of these two approximations is equal to the exact area.

## 2-7 Finding Areas by Antidifferentiation

Suppose we are given the graph of a continuous nonnegative function, represented by the equation $y = f(x)$. Let it be required to find the area $A$ between the graph and the $x$-axis, from $x = a$ to $x = b$, where $a < b$.

Instead of thinking exclusively about the area just described, let us think about the part of the area bounded on the left by the line $x = a$ and on the right by an arbitrary ordinate to the curve (see Fig. 2-26). Let $S$ be this partial area. It depends on the abscissa $x$ of the ordinate at the right. Hence $S$ is a function of $x$, which we denote by writing $S = F(x)$. Evidently $F(a) = 0$ and $F(b)$ is the area $A$ which we are trying to find.

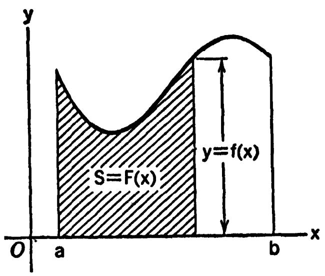



**Fig. 2-26**

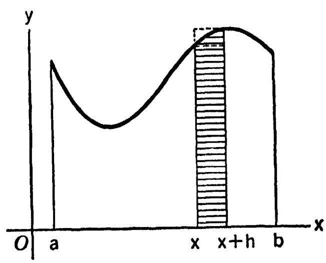



**Fig. 2-27**

As $x$ increases from $a$ to $b$, $S$ increases from 0 to $A$. Let us inquire *how fast* $S$ increases with respect to $x$; that is, let us try to find the value of the derivative $dS/dx$. We know that, by definition,

$$\frac{dS}{dx} = \lim_{h \to 0} \frac{F(x + h) - F(x)}{h}.$$

Let us consider the representation of $F(x + h) - F(x)$ on a diagram. Figure 2-27 shows the situation if $h > 0$. $F(x + h)$ is the area between the ordinates at $a$ and $x + h$; so $F(x + h) - F(x)$ is the shaded area between the ordinates at $x$ and $x + h$. In Fig. 2-27 this area is intermediate in value between the area of a rectangle of height $f(x)$ and one of height $f(x + h)$. That is,

$$hf(x) < F(x + h) - F(x) < hf(x + h).$$

Consequently
$$f(x) < \frac{F(x + h) - F(x)}{h} < f(x + h). \tag{1}$$

If we let $h$ approach 0, the inequalities (1) show that

$$\lim_{h \to 0} \frac{F(x + h) - F(x)}{h} = f(x), \tag{2}$$

because $f$ is continuous, and $f(x + h) \to f(x)$. If $h < 0$, so that the ordinate for $x + h$ is on the left of the one for $x$, the inequalities in (1) turn out to be reversed, but we still arrive at (2) when $h \to 0$. Hence, at least for the case shown in Fig. 2-27, we see that

$$\frac{dS}{dx} = f(x). \tag{3}$$

This is a most important result. *The rate of change of $S$ with respect to $x$ is equal to the length of the ordinate $f(x)$ at the abscissa $x$ (that is, the ordinate at the right edge of the area $S$).*

If the curve had a different appearance from that in Fig. 2-26, the details of the foregoing argument might be slightly different, but the fundamental result (3) would still be obtained. The reasoning, if carried out generally, uses only the fact that the function $f$ is continuous, and not the special appearance of the graph. These matters are considered more thoroughly when we study integrals in a later chapter.

The use of (3) will now be illustrated.

**Example:** Find the area between the curve $y = \frac{1}{4}x^2$ and the $x$-axis, from $x = 0$ to $x = 4$.

This area was found in § 2-6, by the limiting procedure using minor and major approximations by rectangles. We now use the method of antidifferentiation. In this case $f(x) = \frac{1}{4}x^2$, $a = 0$, $b = 4$. Hence we have

$$\frac{dS}{dx} = \frac{1}{4}x^2.$$

Solving this differential equation, we have

$$S = \frac{1}{12}x^3 + C,$$

where $C$ is some constant. Now $S = 0$ when $x = 0$. Hence $0 = 0 + C$, or $C = 0$. Therefore

$$S = \frac{1}{12}x^3.$$

This is the area from 0 out to a variable $x$. When $x = 4$ we obtain the desired area $A$:

$$A = \frac{1}{12}(4)^3 = \frac{16}{3}.$$

This result agrees with that found in § 2-6.

The procedure in general may be described as follows: *Find the general solution of the differential equation* (3). *Put $x = a$ and $S = 0$ to find the value of the constant $C$. Then put $x = b$, and the resulting value of $S$ is the required area $A$.*

## EXERCISES

**1.** In each part of the exercise a straight line and two values of $x$ are given. Find the area between the line and the $x$-axis, from one value of $x$ to the other, by the methods of this section. Check answers by using formulas from elementary geometry.
(a) $y = x + 1$, $x = -1, 2$.
(b) $2x + 3y = 9$, $x = -\frac{3}{2}, \frac{9}{2}$.
(c) $2y = x + 4$, $x = 1, 6$.
(d) $y = 2x + 4$, $x = -1, 3$.
(e) $3x + 2y = 6$, $x = -4, 1$.
(f) $2x + 5y = 20$, $x = -2, 5$.

**2.** Find the area between each parabola and the $x$-axis. Draw a figure in each case.
(a) $x^2 + y = 4$.
(b) $x^2 + 8y = 16$.
(c) $5x^2 + 9y = 45$.
(d) $x^2 + 9y = 36$.
(e) $4x^2 + 9y = 24x$.
(f) $x^2 + 6x + 3y = 0$.
(g) $x^2 - 10x + 4y + 9 = 0$.
(h) $x^2 + 2x + 6y = 8$.

**3.** Find the area between the indicated parabola and the $x$-axis, from $x = a$ to $x = b$, with $a$ and $b$ as given. Draw a figure in each case.
(a) $y = \frac{1}{3}x^2 + 2$, $a = -3$, $b = 3$.
(b) $2y = 3x^2 + 1$, $a = -1$, $b = 3$.
(c) $y = x^2$, $a = -2$, $b = -1$.
(d) $y = x^2 - 2x + 2$, $a = 0$, $b = 2$.
(e) $4y = x^2 - 6x + 17$, $a = 1$, $b = 4$.
(f) $y = x^2 + 4x + 6$, $a = -5$, $b = 0$.

**4.** If a parabola with vertex at the origin and axis along the $y$-axis goes through the point $(B, H)$, show that its equation is $B^2y = Hx^2$. Draw the graph, assuming $B$ and $H > 0$. Show that the area between the curve and the $x$-axis from $x = 0$ to $x = B$ is $BH/3$.

**5.** A parabolic arch is formed by the part of curve $4Hx^2 + B^2y = B^2H$ above the $x$-axis ($B$ and $H > 0$). Show that the area of the arch opening is $\frac{2}{3}BH$, which is just $\frac{2}{3}$ the area of a rectangular opening of the same altitude and base width.

**6.** In each case the curve has just one "arch" above the $x$-axis. Find where the curve crosses the $x$-axis, locate the arch, and then find its area.
(a) $y = -x^3 + 9x$. (d) $y = x^3 - 2x^2 - 5x + 6$.
(b) $y = x^3 - 4x$. (e) $y = x^4 - 10x^2 + 9$.
(c) $y = 5 - \frac{27}{2}x + \frac{15}{2}x^2 - x^3$. (f) $y = -x^4 + 3x^3 + x - 3$.

**7.** Find the area between the curve $y = 8x^3 - 48x^2 + 72x$ (see Fig. 1-32) and the $x$-axis:
(a) from $x = 0$ to $x = 3$;
(b) from $x = 1$ to $x = 3$;
(c) from $x = 3$ to $x = 4$.

**8.** Find the area between the curve $y = -x^4 + 4x^3 + 2x^2 - 12x + 11$ and the $x$-axis:
(a) from $x = -1$ to $x = 3$;
(b) from $x = 0$ to $x = 2$;
(c) from $x = 2$ to $x = 3$.
The graph in this case is like that of Fig. 1-34, but with the positive direction of the $y$-axis reversed.

**9.** Let $S$ be the area between the curve $y = 8x^3 - 48x^2 + 72x$ and the $x$-axis, from $x = 0$ to a variable $x$, where $0 \le x \le 3$. See Fig. 1-32. Find $S$ as a function of $x$ and draw the graph of the function. (a) What is the rate of change of $S$ with respect to $x$ at $x = 0, 1, 2, 3$? (b) For what value of $x$ is $S$ increasing most rapidly? (c) Find the value of $x$ for which $S$ is $\frac{5}{8}$ of the total area from $x = 0$ to $x = 3$.

**10.** (a) Write out the details of the derivation of (3) for a case in which the values of $f$ decrease as the abscissa increases from $x$ to $x + h$. Draw an appropriate figure. (b) Let $f$ be defined as follows: $f(x) = \frac{3}{2}x$ if $0 \le x \le 2$; $f(x) = 3$ if $2 \le x \le 3$; $f(x) = 6 - x$ if $3 \le x \le 6$. Let $S = F(x)$ be defined as in Fig. 2-26 for this case, with $a = 0$, $b = 6$. Compute $F(x)$ by elementary geometry for each of the three intervals $0 \le x \le 2$, $2 \le x \le 3$, $3 \le x \le 6$, and verify in each case that $F'(x) = f(x)$.

## Review Questions and Problems for Chapters I and II

### CONCEPTS AND DEFINITIONS

**1.** What is the absolute value of a number?

**2.** Define the slope of a straight line.

**3.** If $P$ is the point $(x, y)$ on the circle of radius $r$ with center at $O$, and if $OP$ makes the angle $\theta$ with the positive $x$-axis, define the six trigonometric functions of $\theta$.

**4.** What is the definition of the graph of an equation in $x$ and $y$, where $(x, y)$ symbolizes a point in the $xy$-plane?

5. What is a function (single-valued)? If $f$ is such a function, state one important property of the graph of $y = f(x)$ which is not necessarily a property of the graph of an equation (e.g., not a property of the graph of $x^2 + y^2 = 16$).

6. Give an exact definition of the derivative of the function $f$ at the value $x_0$ of the independent variable, using functional notation.

7. Explain the relation between the derivative concept and the concept of instantaneous velocity; of instantaneous acceleration.

8. Define the tangent line to the graph of $y = f(x)$, at a point on it. Explain the relation between the tangent and the derivative.

9. Define the normal to a curve at a point.

10. What is a polynomial? What is a rational function?

11. What is meant by saying "the function $f$ is continuous at $x_0$"? Can you give an example of a function that is not continuous at some point?

12. Explain clearly what is meant by the statement $\lim_{x \to 5} f(x) = 11$.

13. What is meant by an antiderivative of a function? Can a function have more than one antiderivative? Explain.

14. What is a parabola?

15. What is the basic principle used in defining the area of a curved plane figure? Explain how this is applied to the case of the area between the $x$-axis and the curve $y = f(x)$ from $x = a$ to $x = b$, where $f$ is a positive continuous function.

THEORY

1. What theorem of Euclidean geometry is embodied in the formula for the distance between $(x_1, y_1)$ and $(x_2, y_2)$?

2. Two lines with slopes $m_1$, $m_2$ are perpendicular if and only if $m_1m_2 = -1$. What is the trigonometric explanation of this assertion?

3. What is the significant general statement about the form of the equation of every straight line?

4. Derive (i.e., develop by logical steps, starting from the definition of the derivative) the formula for $dy/dx$ if $y = x^n$, where $n$ is a positive integer.

5. Prove that if $f$ is differentiable at $x_0$, it is continuous there.

6. State three basic theorems about limits.

7. What can you say about a differentiable function $f$ on an interval where $f'(x) > 0$? Justify your answer by the law of the mean.

8. What is the relationship between roots and factors of polynomials? Can you demonstrate this relationship?

9. If a differentiable function $f$ is such that $f(x) \leq f(2)$ when $1 \leq x \leq 3$, show by the definition of the derivative that $f'(2) = 0$.

**10.** State the meaning of the law of the mean in geometrical language, and illustrate with a graph. Write the formula, accompanied by a brief statement, which conveys the full substance of the law of the mean.

**11.** Prove that if $f$ is a differentiable function for which $f'(x) = 0$ on an interval, then $f(x_1) = f(x_2)$ for every choice of $x_1$, $x_2$ on the interval. Of what significance is this fact in relation to the finding of antiderivatives?

**12.** What general statement can be made about the type of curve represented by an equation of one of the forms $y = Ax^2 + Bx + C$, $x = Ay^2 + By + C$, where $A \neq 0$?

## PROBLEMS

**1.** Let $P_1$ be the point $(3, \frac{9}{4})$ on the parabola $4y = x^2$. A line is drawn through $P_1$ and the focus; it intersects the parabola again at a point $P_2$. Find the equations of the tangents to the parabola at $P_1$ and $P_2$, respectively. Show that they intersect at right angles at a point on the directrix. Find this point.

**2.** Solve the preceding problem if $P_1$ is any point $(x_1, y_1)$ on the parabola, with $x_1 > 0$. In this case $P_2$ is the point $(-4/x_1, 4/x_1^2)$, and the tangents intersect at $\left(\dfrac{x_1^2 - 4}{2x_1}, -1\right)$.

**3.** Show that the parabolas $2py + p^2 = x^2$, $2qy + x^2 = q^2$, where $p$ and $q$ are positive, have the same focus. What point is it? Draw a few curves of each type. Prove that each curve of one type intersects every curve of the other type orthogonally.

**4.** A line is drawn from the origin tangent to the parabola $2y - 4 = -(x - 4)^2$ at a point in the first quadrant. Find this point. What is the other point of the parabola where a line through the origin is tangent to the curve?

**5.** The height of a cone is 18 inches. It remains constant while the radius of the base increases $\frac{1}{4}$ inch per second, starting from a radius of 5 inches at $t = 0$. Find the rate of increase of the volume of the cone when the base radius is 10 inches.

**6.** A cube and a sphere are increasing in size. The edge of the cube and the radius of the sphere are 5 and 6 inches, respectively, at $t = 0$, and they increase at the rates of 4 and 3 inches per minute, respectively. Find the dimensions of the cube and sphere when the volume of the sphere is increasing $\pi$ times as fast as the volume of the cube.

**7.** A trough whose cross section is an isosceles triangle is 6 feet long, 2 feet across the top, and 18 inches deep. If water in the trough is initially 6 inches deep and increases in depth at $\frac{1}{8}$ inch per second, how fast is the volume of water in the trough increasing when the trough is $\frac{4}{9}$ filled?

**8.** Draw the graph of $y = x^4 - 12x^3 + 28x^2 - 20$. What inequalities must $A$ satisfy if the equation $x^4 - 12x^3 + 28x^2 + A = 0$ is to have 4 distinct real roots? For what values of $A$ does the equation have no real roots?

9. Draw the graph of $y = x^4 - 6x^2 + 8x + 10$. Is there any value of $A$ such that the equation $x^4 - 6x^2 + 8x + A = 0$ has 4 distinct real roots? For what values of $A$ does it have no real roots? What is the situation about real roots if $A = 24$? If $A = -3$?

10. Prove that the curves $y = x^3 - 3x^2 - 8x - 4$, $y = 3x^2 + 7x + 4$ have a common tangent at just one point. Find the point. Draw both curves.

11. A normal is drawn to the parabola $x^2 = 2py$ $(p > 0)$ at a point $P$ in the first quadrant. Let $G$ be the intersection of this normal with the $y$-axis, and let $N$ be the foot of the perpendicular from $P$ to the $y$-axis. Show that the length $NG$ is constant as $P$ varies.

12. A right circular cylinder of height $2x$ is inscribed in a sphere of radius 5 units. Express the volume $y$ of the cylinder as a function of $x$ (it is a cubic polynomial). Draw the graph of this polynomial. What values of $x$ are of significance in this problem? From the graph read off the maximum possible $y$ and the value of $x$ for which it occurs.

13. Find $y$ in terms of $x$ if:
    (a) $dy/dx = 8x^3 - 2x$ and $y = 8$ when $x = 2$.
    (b) $dy/dx = x^2 - x$ and $y = 1$ when $x = 3$.

14. Find the area enclosed between the parabola $y = 6 + x - x^2$ and the $x$-axis.

15. Prove by mathematical induction that the sum of the first $n$ odd positive integers is $n^2$, i.e., that $1 + 3 + 5 + \cdots + (2n - 1) = n^2$.

16. A point moves on the $s$-axis with acceleration $-32 - 16t$ feet per second per second. Find $s$ in terms of $t$ ($s$ in feet, $t$ in seconds), given that $s = 0$ and $v = ds/dt = 40$ when $t = 0$. Draw the graph. During what intervals of time is $s$ increasing as $t$ increases?

17. A ball is rolling up an incline with a negative acceleration $-18$ feet per second per second. If $s$ is measured up the incline, if $s = 44$ when $t = 2$, and if $s = 71$ when $t = 5$, find $s$ in terms of $t$. What is the farthest point up the incline reached by the ball?

18. A car goes through a red traffic light at 30 miles per hour, and continues along a straight highway at the same speed. A police car at the light starts in pursuit 2 seconds later, with initial velocity 0 and constant acceleration 6 feet per second per second. When and where will the police car overtake the offender, and at what ultimate speed will the police car be going?

*CHAPTER* III

# DIFFERENTIATION OF ALGEBRAIC FUNCTIONS

## 3-1 The Δ-Notation

As we already know, the derivative with respect to $x$ of a function $f$ is defined as the limit

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

If we write $y = f(x)$, it is customary and convenient to use what is known as the Δ-notation to express changes in the value of $y$ corresponding to changes in the value of $x$, and this notation is commonly used in expressing the definition of the derivative. Let $x$ and $x + \Delta x$ denote two values of the independent variable. The symbol $\Delta x$ (read delta-$x$) represents the *change* in the independent variable $x$. In general we use the symbol Δ (Greek capital letter delta) as a prefix to indicate a change in the value of the letter variable which follows the Δ. The Δ by itself is not a number, but $\Delta x$ is a symbol to which we can assign numerical values. The value of the function $f$, when $x$ is replaced by $x + \Delta x$, is denoted by $y + \Delta y$, so that

$$y = f(x), \qquad y + \Delta y = f(x + \Delta x),$$

and so

$$\Delta y = f(x + \Delta x) - f(x)$$

is the change in $y$ corresponding to the change in $x$ represented by $\Delta x$.

The definition of the derivative can now be written in either of the forms

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{1}$$

or
$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}. \tag{2}$$

All of this involves no new fundamental concepts; we are just explaining a conventional notation.

**Example 1:** Find $\Delta y$ in terms of $x$ and $\Delta x$ if $y = 1/x^2$. Express $\Delta y$ as a single fraction and simplify it. Then calculate $dy/dx$.

We write

$$y + \Delta y = \frac{1}{(x + \Delta x)^2}, \qquad y = \frac{1}{x^2},$$

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

$$\Delta y = \frac{-2x\,\Delta x - (\Delta x)^2}{(x + \Delta x)^2 x^2}.$$

Consequently
$$\frac{\Delta y}{\Delta x} = \frac{-2x - \Delta x}{(x + \Delta x)^2 x^2},$$

and
$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{-2x}{x^4} = \frac{-2}{x^3}.$$

Here we have used Theorems 1-C, 1-D, and 1-E from § 1-8.

**Example 2:** If $R$ is the ratio of the ages of a father and his son, who is 30 years younger, find the change in $R$ corresponding to a change in the age of the father, whose age is $x$ years. Then find $dR/dx$ when $x = 35$, and also when $x = 60$.

We have

$$R = \frac{x}{x - 30}, \qquad R + \Delta R = \frac{x + \Delta x}{(x + \Delta x) - 30},$$

$$\Delta R = \frac{x + \Delta x}{x + \Delta x - 30} - \frac{x}{x - 30} = \frac{(x + \Delta x)(x - 30) - x(x + \Delta x - 30)}{(x + \Delta x - 30)(x - 30)}.$$

On simplification we find

$$\Delta R = \frac{-30\,\Delta x}{(x + \Delta x - 30)(x - 30)}, \qquad \frac{\Delta R}{\Delta x} = \frac{-30}{(x + \Delta x - 30)(x - 30)}.$$

Then
$$\frac{dR}{dx} = \lim_{\Delta x \to 0} \frac{\Delta R}{\Delta x} = \frac{-30}{(x - 30)^2}.$$

If $x = 35$ or $x = 60$ we have, respectively,

$$\frac{dR}{dx} = -\frac{6}{5} \quad \text{or} \quad \frac{dR}{dx} = -\frac{1}{30}.$$

Evidently the ratio $R$ changes more and more slowly with respect to $x$ as $x$ gets larger and larger.

## 3-2 Sums, Products, and Quotients

In this section we derive some general rules which will help us to calculate the derivatives of a great variety of functions. We state these rules as theorems. Recall that a function is called differentiable if it has a derivative.

THEOREM 3-A. *If u and v are differentiable functions of x, then u + v is also differentiable, and*

$$\frac{d}{dx}(u + v) = \frac{du}{dx} + \frac{dv}{dx}.$$

*In other words, the derivative of a sum is the sum of the derivatives. This also applies to sums of more than two functions:*

$$\frac{d}{dx}(u_1 + u_2 + \cdots + u_n) = \frac{du_1}{dx} + \cdots + \frac{du_n}{dx},$$

*where n is any positive integer and* $u_1, \cdots, u_n$ *are differentiable functions of x.*

*Proof.* Let $y = u + v$. For a change $\Delta x$ in the independent variable we have

$$y + \Delta y = (u + \Delta u) + (v + \Delta v),$$

and so

$$\Delta y = \Delta u + \Delta v.$$

Therefore

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}.$$

It follows by Theorem 1-C, when we take limits as $\Delta x \to 0$, that

$$\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}.$$

The proof extends at once to the case of $n$ functions, by the use of mathematical induction.

The next theorem concerns the handling of constant factors.

THEOREM 3-B. *If u is a differentiable function of x, and c is a constant, cu is also differentiable, and*

$$\frac{d}{dx}(cu) = c\frac{du}{dx}.$$

*That is, the derivative of a constant multiple of a function is equal to that constant multiplied by the derivative of the function.*

*Proof.* Let $y = cu$. For a change $\Delta x$ in the independent variable we have

$$y + \Delta y = c(u + \Delta u),$$

and so

$$\Delta y = c(u + \Delta u) - cu = c\,\Delta u,$$

whence

$$\frac{\Delta y}{\Delta x} = c\frac{\Delta u}{\Delta x}.$$

Then, on taking limits as $\Delta x \to 0$, we have

$$\frac{dy}{dx} = c\frac{du}{dx}.$$

In particular, if $c = -1$, we have

$$\frac{d}{dx}(-u) = -\frac{du}{dx}.$$

Therefore, on combining this result with Theorem 3-A, we see that

$$\frac{d}{dx}(u - v) = \frac{du}{dx} - \frac{dv}{dx}.$$

We have already been using Theorems 3-A and 3-B in a special case in the differentiation of polynomials, e.g.,

$$\frac{d}{dx}(3 - 4x + 7x^2) = 0 - 4\cdot 1 + 7\cdot 2x = -4 + 14x.$$

The rules for differentiating products and quotients come next.

THEOREM 3-C. *If $u$ and $v$ are differentiable functions of $x$, $uv$ is also differentiable, and*

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}. \tag{1}$$

*In words we render this rule as follows: The derivative of the product of two functions is*

*first times derivative of second*

*plus*

*second times derivative of first.*

*Proof.* Let $y = uv$. For a change $\Delta x$ in the independent variable we have

$$\begin{aligned} y + \Delta y &= (u + \Delta u)(v + \Delta v) \\ &= uv + v\,\Delta u + u\,\Delta v + (\Delta u)(\Delta v), \end{aligned}$$

and so

$$\Delta y = v\,\Delta u + u\,\Delta v + (\Delta u)(\Delta v),$$

$$\frac{\Delta y}{\Delta x} = v\frac{\Delta u}{\Delta x} + u\frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta x}\Delta v.$$

Then, by Theorems 1-C and 1-D,

$$\lim \frac{\Delta y}{\Delta x} = v \lim \frac{\Delta u}{\Delta x} + u \lim \frac{\Delta v}{\Delta x} + \lim \frac{\Delta u}{\Delta x} \lim \Delta v.$$

For convenience in printing we have omitted the symbols $\Delta x \to 0$ under the limit abbreviation lim. It now follows that

$$\frac{dy}{dx} = v\frac{du}{dx} + u\frac{dv}{dx} + \frac{du}{dx}\cdot 0,$$

and this is equivalent to the formula for $(d/dx)(uv)$ given in the theorem. We use the fact that $\Delta v \to 0$ when $\Delta x \to 0$. This expresses the continuity of $v$, which is a consequence of the differentiability (Theorem 1-F, § 1-8).

The rule of Theorem 3-C can be applied in stages to the product of three or more factors:

$$\begin{aligned}\frac{d}{dx}(uvw) &= u\frac{d}{dx}(vw) + vw\frac{du}{dx} \\ &= u\left(v\frac{dw}{dx} + w\frac{dv}{dx}\right) + vw\frac{du}{dx} \\ &= uv\frac{dw}{dx} + uw\frac{dv}{dx} + vw\frac{du}{dx}.\end{aligned}$$

A more symmetrical way of writing this is

$$\frac{d}{dx}(u_1u_2u_3) = \frac{du_1}{dx}u_2u_3 + u_1\frac{du_2}{dx}u_3 + u_1u_2\frac{du_3}{dx}.$$

The rule extends in like manner to the product of more than three functions.

**Example 1:** Find $dy/dx$ from $y = (x^2 - 4)(x^3 - 8)$.
The rule gives

$$\begin{aligned}\frac{dy}{dx} &= (x^2 - 4)\frac{d}{dx}(x^3 - 8) + (x^3 - 8)\frac{d}{dx}(x^2 - 4) \\ &= (x^2 - 4)\cdot 3x^2 + (x^3 - 8)\cdot 2x.\end{aligned}$$

We can now pick out $x$ and $x - 2$ as common factors.

$$\begin{aligned}\frac{dy}{dx} &= x(x - 2)[3x(x + 2) + 2(x^2 + 2x + 4)] \\ &= x(x - 2)(5x^2 + 10x + 8).\end{aligned}$$

THEOREM 3-D. *If $u$ and $v$ are differentiable functions of $x$, then $u/v$ is also differentiable whenever $v \neq 0$, and then*

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\dfrac{du}{dx} - u\dfrac{dv}{dx}}{v^2}. \tag{2}$$

*That is, the derivative of a quotient follows the rule*

$$\frac{d}{dx}\text{(quotient)} = \frac{\text{(denominator)}\dfrac{d}{dx}\text{(numerator)} - \text{(numerator)}\dfrac{d}{dx}\text{(denominator)}}{\text{(denominator)}^2}.$$

*Proof.* Let $y = u/v$. For a change $\Delta x$ in the independent variable we have

$$y + \Delta y = \frac{u + \Delta u}{v + \Delta v}, \qquad \Delta y = \frac{u + \Delta u}{v + \Delta v} - \frac{u}{v},$$

$$\Delta y = \frac{uv + v\,\Delta u - uv - u\,\Delta v}{(v + \Delta v)v} = \frac{v\,\Delta u - u\,\Delta v}{(v + \Delta v)v}.$$

Then
$$\frac{\Delta y}{\Delta x} = \frac{v\dfrac{\Delta u}{\Delta x} - u\dfrac{\Delta v}{\Delta x}}{(v + \Delta v)v}.$$

When $\Delta x \to 0$ it follows that $\Delta v \to 0$, because $v$ is differentiable, and therefore continuous. Hence, using Theorems 1-C, 1-D, and 1-E, we obtain the result

$$\frac{dy}{dx} = \frac{v\dfrac{du}{dx} - u\dfrac{dv}{dx}}{v^2}.$$

A particular case which is frequently used is that in which $u$ is constant: $u = c$. Then $du/dx = 0$, and so we have

$$\frac{d}{dx}\left(\frac{c}{v}\right) = -\frac{c}{v^2}\frac{dv}{dx}. \tag{3}$$

**Example 2:** Find $dy/dx$ if $y = (3x - 5)/(x^2 + 9)$.

$$\frac{dy}{dx} = \frac{(x^2 + 9)\dfrac{d}{dx}(3x - 5) - (3x - 5)\dfrac{d}{dx}(x^2 + 9)}{(x^2 + 9)^2}$$
$$= \frac{(x^2 + 9)\cdot 3 - (3x - 5)\cdot 2x}{(x^2 + 9)^2}$$
$$= \frac{3x^2 + 27 - 6x^2 + 10x}{(x^2 + 9)^2} = \frac{-3x^2 + 10x + 27}{(x^2 + 9)^2}.$$

**Example 3:** Find $dy/dx$ if $y = 8/(4 - x^3)$. By (3) we have

$$\frac{dy}{dx} = \frac{-8}{(4 - x^3)^2}\frac{d}{dx}(4 - x^3) = \frac{-8(-3x^2)}{(4 - x^3)^2},$$
$$\frac{dy}{dx} = \frac{24x^2}{(4 - x^3)^2}.$$

**Example 4:** We shall prove that the exponent rule

$$\frac{d}{dx}x^n = nx^{n-1} \tag{4}$$

holds true if $n$ is a *negative* integer.

Let $y = x^n$. Then $-n = p$ is a *positive* integer, and $y = 1/x^p$. Therefore

$$\frac{dy}{dx} = \frac{d}{dx}\frac{1}{x^p} = \frac{-1}{(x^p)^2}\frac{d}{dx}(x^p) = -\frac{px^{p-1}}{x^{2p}},$$

or
$$\frac{dy}{dx} = -px^{p-1-2p} = -px^{-p-1} = nx^{n-1}.$$

We now know that the exponent rule (4) is valid for both positive and negative *integers*. Actually, it is valid for all constant values of $n$, but we are not yet ready to prove this general assertion of the rule.

## EXERCISES

Find $dy/dx$ in each part of Exercises 1–3. Simplify your answers by factoring whenever possible.

**1.** (a) $y = (3x^2 + 1)(x^3 + 6x)$.
(b) $y = (x^2 - 3x)(8 - x^3)$.
(c) $y = (x^2 - 1)(x^3 - 1)$.
(d) $y = (x + 1)(x^2 - 4x - 5)$.
(e) $y = (x^2 - 4x + 3)(x^2 - 2x - 3)$.
(f) $y = (x^2 - 1)(x^3 - 2x - 1)$.
(g) $y = (x^4 - x^2)(x^3 - 1)$.
(h) $y = x(x + 1)(x + 2)$.
(i) $y = x^2(x^2 - 4)(x^3 + 8)$.
(j) $y = (x^2 - 9)(x^3 - 27)(x^4 - 81)$.
(k) $y = x^4(4 - x^2)(16 - x^4)$.
(l) $y = (3x + 2)(9x^2 - 4)(27x^3 + 8)$.

**2.** (a) $y = \dfrac{1 - x}{1 + x}$.
(b) $y = \dfrac{1 - x^2}{1 + x^2}$.
(c) $y = \dfrac{2x + 1}{x^2 + 2}$.
(d) $y = \dfrac{x^2}{3 - x}$.
(e) $y = \dfrac{x^3}{1 - x^2}$.
(f) $y = \dfrac{8x}{25 - x^2}$.
(g) $y = \dfrac{2x + 1}{x^2 + x - 4}$.
(h) $y = \dfrac{x^2 + 6x + 9}{x^2 - 4x + 4}$.
(i) $y = \dfrac{x^2 - x}{x^2 - 4}$.
(j) $y = \dfrac{x^2}{(x + 1)(x - 2)}$.
(k) $y = \dfrac{2x}{(x^2 - 4)(x^3 - 8)}$.
(l) $y = x^2\dfrac{a + x}{a - x}$.
(m) $y = \dfrac{2ax^3 - x^4}{x - a}$.
(n) $y = \dfrac{8x}{x^4 - 8x^2 + 16}$.

**3.** (a) $y = \dfrac{18}{x^2 - 9}$.
(b) $y = \dfrac{3}{2x - 5}$.
(c) $y = \dfrac{1}{8 - x^3}$.
(d) $y = \dfrac{50}{x^2 + 25}$.
(e) $y = \dfrac{x}{4x^2 + 5}$.
(f) $y = \dfrac{4x - x^4}{x^3 + 2}$.
(g) $y = \dfrac{x^3 - 4x}{16 - x^2}$.
(h) $y = \dfrac{x^2 - 4x + 2}{x - 4}$.
(i) $y = \dfrac{x^2 + 5x - 7}{2x - 3}$.
(j) $y = \dfrac{x^5}{x^2 + 1}$.
(k) $y = \dfrac{2x^2}{x^2 + 4}$.
(l) $y = \dfrac{x^3}{2a - x}$.

4. Find $dy/dx$ in two ways in each case: once using the exponent rule, and once using the rule for quotients. Show that your answers by the two methods agree.

(a) $y = 3x^{-4} - 4x^{-3}$.

(b) $y = x^{-4} + 6x^{-2} - 60$.

(c) $y = 2x^{-3} - 3x^{-2} - 36x^{-1} + 20$.

(d) $y = \dfrac{1 - 4x}{x}$.

(e) $y = \dfrac{9 - x^3}{x^2}$.

(f) $y = \dfrac{10 + 30x^4 - 10x^6}{x^5}$.

(g) $y = \dfrac{3}{x^3} - \dfrac{2}{x^2} - \dfrac{5}{x} + 3$.

(h) $y = \dfrac{8(x + 1)}{x}$.

(i) $y = \dfrac{x^3 - 12x}{x^4}$.

(j) $y = \dfrac{7 - 4x}{x^4} + \dfrac{5 - 9x + x^2}{x^2}$.

## 3-3 Composite Functions

We frequently construct functions by putting the value of one function in place of the independent variable in another function. Functions constructed in this manner are called *composite functions.*

**Example 1:**

(a) If in $y = u^2$ we substitute $u = 5x^2 - x$, we obtain $y = (5x^2 - x)^2$.

(b) If in $y = \sqrt{u}$ we substitute $u = \dfrac{x + 1}{x - 1}$, we obtain $y = \sqrt{\dfrac{x + 1}{x - 1}}$.

**Example 2:**

(a) $y = \dfrac{1}{1 + u^2}$, $u = 2x - 3$, $y = \dfrac{1}{1 + (2x - 3)^2}$.

(b) $y = u^3$, $u = \log_{10} x$, $y = (\log_{10} x)^3$.

Suppose $\phi$ and $f$ are functions of the independent variables $x$ and $u$, respectively, and suppose that the values $\phi(x)$ are in the domain of definition of the function $f$. If we then set $u = \phi(x)$ and write

$$y = f(u) = f[\phi(x)],$$

this notation expresses the formation of a composite function in which $x$ is the independent variable and $y$ is the dependent variable. The intermediate variable $u$ has two roles: in $y = f(u)$ it is independent and in $u = \phi(x)$ it is dependent. We shall now see that if we know the two derivatives $du/dx$ (with $x$ independent) and $dy/du$ (with $u$ independent), we can calculate the derivative $dy/dx$ (with $x$ independent).

THEOREM 3-E. *Let $\phi$ and $f$ be differentiable functions of $x$ and $u$, respectively, and let the composite function be denoted by $F$:*

$$y = f(u), \qquad u = \phi(x),$$
$$y = F(x) = f[\phi(x)].$$

*Then $F$ is a differentiable function of $x$, and the derivative of $y$ with respect to $x$ is*

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}, \tag{1}$$

*or, alternatively,*

$$F'(x) = f'[\phi(x)] \cdot \phi'(x) = f'(u) \cdot \phi'(x). \tag{2}$$

We call this rule (1) or (2) the *composite function rule.* It is also called the *chain rule.* The use of this rule is very important and convenient in practice. For instance, it enables us to avoid lengthy expansions by the binomial theorem.

**Example 3:** If $y = (3x^2 - 2x + 1)^4$, find $dy/dx$.

We *could* obtain the answer by working out the indicated fourth power, using the binomial theorem. We could then differentiate the resulting eighth-degree polynomial. But it is vastly simpler to write

$$y = u^4, \qquad u = 3x^2 - 2x + 1,$$

and use the composite function rule:

$$\frac{dy}{du} = 4u^3, \qquad \frac{du}{dx} = 6x - 2,$$

$$\frac{dy}{dx} = 4(3x^2 - 2x + 1)^3 \cdot (6x - 2)$$

$$= 8(3x^2 - 2x + 1)^3(3x - 1).$$

Now let us consider the proof of Theorem 3-E. With the regular $\Delta$-notation we have

$$\lim_{\Delta u \to 0} \frac{\Delta y}{\Delta u} = \frac{dy}{du}, \qquad \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} = \frac{du}{dx}.$$

At a first glance it appears as though we could make the proof as follows:

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}, \tag{3}$$

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta u \to 0} \frac{\Delta y}{\Delta u} \cdot \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

But this argument has a defect. In order to be able to write (3) we need to have $\Delta x \neq 0$ and $\Delta u \neq 0$. Now, in the definition of $dy/dx$ as the limit of $\Delta y/\Delta x$, $\Delta x$ is the independent variable, and it is required that $\Delta x \neq 0$. But there is no guarantee that $\Delta u \neq 0$. It can in fact happen that $\Delta u = 0$, because $\phi(x)$ and $\phi(x + \Delta x)$ may have the same value, even when $\Delta x$ is very small, and the definition of $\Delta u$ is

$$\Delta u = \phi(x + \Delta x) - \phi(x). \tag{4}$$

To get around the defect in the argument based on (3), we start out with $\Delta x$ as the independent variable; we define $\Delta u$ by (4) and then we define $\Delta y$ as follows:

$$\Delta y = f(u + \Delta u) - f(u). \tag{5}$$

Observe that $\Delta y = 0$ if $\Delta u = 0$. Observe also from (4) that $\Delta u \to 0$ if $\Delta x \to 0$; this is because $\phi$ is differentiable, and hence continuous. Now we define $\epsilon$ as a function of $\Delta x$. If $\Delta u$ happens to equal 0, we define $\epsilon = 0$; otherwise we define

$$\epsilon = \frac{\Delta y}{\Delta u} - \frac{dy}{du}. \tag{6}$$

It follows from (6) that

$$\Delta y = \frac{dy}{du}\,\Delta u + \epsilon\,\Delta u,$$

and this formula is still correct if $\Delta u = 0$, for then $\Delta y = 0$ also. From the way in which $\epsilon$ is defined we see that $\epsilon \to 0$ as $\Delta x \to 0$. We now write

$$\frac{\Delta y}{\Delta x} = \frac{dy}{du}\cdot\frac{\Delta u}{\Delta x} + \epsilon\frac{\Delta u}{\Delta x}$$

and let $\Delta x$ approach 0. The result is

$$\frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx} + 0\cdot\frac{du}{dx}.$$

This is the same as (1), so the proof is complete.

One very important application of the chain rule has been illustrated in Example 3. The general principle is that

$$\frac{d}{dx}(\quad)^n = n(\quad)^{n-1}\frac{d}{dx}(\quad),$$

where any differentiable function of $x$ may be inserted in the parentheses. If the function is denoted by $u$, the formula is

$$\frac{d}{dx}\,u^n = nu^{n-1}\frac{du}{dx}. \tag{7}$$

**Example 4:** If $y = \left(\dfrac{1-x}{1+x}\right)^3$,

$$\frac{dy}{dx} = 3\left(\frac{1-x}{1+x}\right)^2\frac{d}{dx}\left(\frac{1-x}{1+x}\right) = 3\left(\frac{1-x}{1+x}\right)^2\frac{(1+x)(-1)-(1-x)}{(1+x)^2}$$
$$= \frac{-6(1-x)^2}{(1+x)^4}.$$

**Example 5:** If $y = (a^2 + x^2)(a^2 - x^2)^{-2}$,

$$\frac{dy}{dx} = (a^2 + x^2)\frac{d}{dx}(a^2 - x^2)^{-2} + (a^2 - x^2)^{-2}\frac{d}{dx}(a^2 + x^2)$$
$$= (a^2 + x^2)(-2)(a^2 - x^2)^{-3}(-2x) + (a^2 - x^2)^{-2}2x.$$

To simplify, we take out the factor $2x$, get rid of the negative exponents, and reduce to a common denominator:

$$\frac{dy}{dx} = 2x\left[\frac{2(a^2 + x^2)}{(a^2 - x^2)^3} + \frac{1}{(a^2 - x^2)^2}\right] = 2x\left[\frac{2a^2 + 2x^2 + a^2 - x^2}{(a^2 - x^2)^3}\right]$$
$$= \frac{2x(x^2 + 3a^2)}{(a^2 - x^2)^3}.$$

It is important to realize that the composite function rule does not depend on the particular letters which are used for the variables. Thus, for instance, if $w = p^5$ and $p$ is a differentiable function of $t$, we have

$$\frac{dw}{dt} = 5p^4\frac{dp}{dt}.$$

And if $z = f(s)$, $s = g(r)$, we have

$$\frac{dz}{dr} = \frac{dz}{ds}\cdot\frac{ds}{dr}.$$

A simple but interesting illustration of the use of the composite function rule is given in the next example.

**Example 6:** A spherical balloon is being inflated. At a certain instant the diameter of the balloon is 3 feet and the diameter is increasing 2 inches per second. Find the rate of increase of the volume of the balloon at that instant.

We denote the volume of the balloon by $V$, its diameter by $D$. The formula for the volume is

$$V = \frac{4}{3}\pi\left(\frac{D}{2}\right)^3 = \frac{\pi}{6}D^3.$$

Now $D$ is increasing in some way with the time $t$. We may think of $D$ as a function of $t$, and so $V$, which is a function of $D$, becomes a composite function of $t$. The time rate of change of $V$ is

$$\frac{dV}{dt} = \frac{dV}{dD}\cdot\frac{dD}{dt} = \frac{\pi}{6}3D^2\frac{dD}{dt} = \frac{\pi}{2}D^2\frac{dD}{dt}.$$

This is a general formula for $dV/dt$, no matter how $D$ changes. For our specific problem we put $D = 3$, $dD/dt = \frac{2}{12} = \frac{1}{6}$ (the foot and the second as units). Then

$$\frac{dV}{dt} = \frac{\pi}{2}\cdot 9\cdot\frac{1}{6} = \frac{3\pi}{4} \quad \text{(cubic feet per second).}$$

## EXERCISES

1. Find $\frac{dy}{dx}$ in each case.

(a) $y = (3 - 2x^3)^2$.
(b) $y = (7x^2 - 5)^3$.
(c) $y = (x^4 - 2x)^5$.
(d) $y = (3x^2 - 4x + 1)^2$.
(e) $y = (36 - x^2)^{-2}$.
(f) $y = (4x^2 + 9)^{-1}$.
(g) $y = (4x^3 - 9x^2)^2(3x - 2x^2)^3$.
(h) $y = (x^2 - 4)^3(x^3 + 8)^2$.
(i) $y = \frac{(16 + x^2)^2}{16 - x^2}$.
(j) $y = \frac{2x - 3x^2}{(1 - 3x)^3}$.

**2.** Find $\frac{dy}{dx}$ in each case. Express answers without negative exponents.

(a) $y = \frac{1}{(3x - 4)^2}$.

(b) $y = \frac{8}{(7 - 2x)^3}$.

(c) $y = \frac{1}{(x^2 - 1)^5}$.

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

(e) $y = \left(\frac{2x + x^2}{x + 1}\right)^{-1}$.

(f) $y = \left(\frac{25 - x^2}{x^2}\right)^{-2}$.

(g) $y = \frac{(2x + 1)^3}{(x^2 - 3)^2}$.

(h) $y = x^2(16 - x^2)^{-2}$.

**3.** Find $\frac{dw}{dv}$ in each case. Express answers without negative exponents.

(a) $w = v^2(v^2 - 25)^3$.

(b) $w = (2v + 3)^3(1 - v^2)$.

(c) $w = \frac{(v + 3v^2)^2}{v + 1}$.

(d) $w = \frac{v^4 - 12v^2}{(v^2 - 4)^2}$.

(e) $w = \frac{(v^2 - 4)^3}{3v^2 + 4}$.

(f) $w = (v^2 - 1)(25 + v^2)^{-3}$.

(g) $w = (1 - 2v)^{-4}(v^2 - v)^2$.

(h) $w = (v - a)^{-1}(2av - v^2)^2$.

**4.** A pebble thrown into a still pond produces a series of concentric circular ripples. If the radius of a ripple is increasing $1\frac{1}{2}$ feet per second, how fast is the area within the ripple increasing when the radius is 10 feet?

**5.** A rectangle of changing dimensions (length $x$ and breadth $y$) has constant area 100 square units. If $x$ increases at the rate of 2 units per minute, find the rate of change of $y$ when $x = 20$ units.

**6.** A horizontal water tank is 6 feet long. Its vertical cross section is an isosceles triangle 2 feet across the top and $1\frac{1}{2}$ feet deep (see Fig. 3-1). Water is being poured into the trough. If the depth of the water is increasing $\frac{1}{2}$ inch per second when the depth is 1 foot, what is the rate of increase of the volume of water in the trough at that instant?

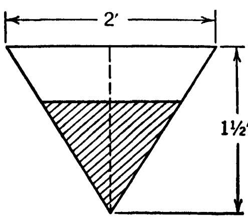


**Fig. 3-1**

**7.** A man 6 feet tall is walking directly away from a post on which there is a lamp 18 feet above the street level. If the man is walking 5 feet per second, how fast is the length of his shadow increasing?

**8.** A flashlight throws a cone of light with a 30° angle between the outermost rays and the axis of the cone. A man points the light straight at a blank wall. How fast is the illuminated area of the wall changing if the light is 9 feet from the wall and is being brought toward the wall at the rate of 6 feet per second?

## 3-4 Second Derivatives

If a function $f$ has a derivative $f'(x)$ at each point of a certain interval, the derivative itself is a function defined on that interval. It may be possible to differentiate $f'(x)$. If so, the derivative of $f'(x)$ is called the *second derivative* of $f(x)$. It is denoted by $f''(x)$. If $y = f(x)$, another notation for the second derivative is

$$\frac{d}{dx}\left(\frac{dy}{dx}\right), \quad \text{or more compactly} \quad \frac{d^2y}{dx^2}.$$

The derivative of $f''(x)$ is called the *third* derivative of $f(x)$; it is denoted by $f'''(x)$ or $f^{(3)}(x)$. Other notations are

$$\frac{d}{dx}\left(\frac{d^2y}{dx^2}\right) = \frac{d^2}{dx^2}\left(\frac{dy}{dx}\right) = \frac{d^3y}{dx^3}.$$

Derivatives of fourth, fifth, and higher orders are defined and denoted analogously. For the derivative of order $n$, where $n$ is any positive integer, we write

$$f^{(n)}(x) \quad \text{or} \quad \frac{d^ny}{dx^n}.$$

It is often convenient, as a saving of space, to write $y'$ instead of $f'(x)$ or $dy/dx$. We also write $y''$ for the second derivative, $y'''$ or $y^{(3)}$ for the third derivative, and so on.

**Example 1:** Find $y''$ and $y'''$ if $y = \dfrac{1+x}{1-x}$. We begin with the first derivative:

$$y' = \frac{(1-x)\cdot 1 - (1+x)(-1)}{(1-x)^2} = \frac{2}{(1-x)^2} = 2(1-x)^{-2}.$$

Then

$$y'' = 2(-2)(1-x)^{-3}(-1) = 4(1-x)^{-3},$$
$$y''' = 4(-3)(1-x)^{-4}(-1) = 12(1-x)^{-4}.$$

*The Direction of Concavity of a Curve*

We shall see, a good deal later in our studies, that there are reasons for wanting to compute derivatives of all the higher orders. Just now we confine our attention mainly to second derivatives, which are useful in studying the graphs of functions. The second derivative enables us to determine the direction of concavity of a curve. In order to explain this we must begin by defining the terms "concave upward" and "concave downward" as applied to the graph of a function.

We consider the graph of $y = f(x)$, where $f$ is a function which is defined and has first and second derivatives at all the points under consideration. A segment of the graph is said to be *concave upward* if for every three points $P_1$, $P_2$, $P_3$ in order from left to right along that portion of the graph, the in-between point $P_2$ is below the line joining the outer points $P_1$ and $P_3$ (see Fig. 3-2). A moment's consideration shows that

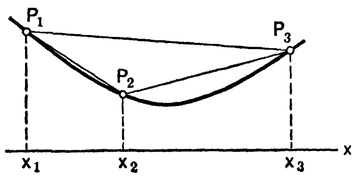



**Fig. 3-2**

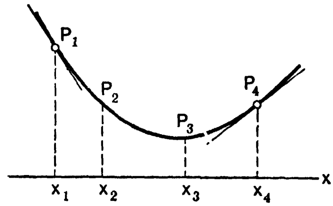



**Fig. 3-3**

this condition is the same as saying that the slope of the line joining two points of the curve must increase as either point moves to the right along the curve. Thus in Fig. 3-2,

$$\text{slope of } P_1P_2 < \text{slope of } P_1P_3 < \text{slope of } P_2P_3.$$

If we consider four points in order, as in Fig. 3-3, we see then that

$$\text{slope of } P_1P_2 < \text{slope of } P_3P_4.$$

But if we let $P_2$ approach $P_1$ and $P_3$ approach $P_4$, the slope of $P_1P_2$ decreases toward the limit $f'(x_1)$ (the slope at $P_1$), and the slope of $P_3P_4$ increases toward the limit $f'(x_4)$. Hence

$$f'(x_1) < f'(x_4).$$

*In other words, $f'(x)$ increases as $x$ increases when the curve is concave upward.* This means, geometrically, that the tangent to the curve turns in the counterclockwise direction as the point of tangency moves to the right.

It is easy to show, conversely, that when $f'(x)$ increases as $x$ increases, the curve is concave upward. For this purpose we use the law of the mean (Theorem 2-C, § 2-1). Suppose $x_1 < x_2 < x_3$. By the law of the mean there are certain values of $x$, say $X_1$ and $X_2$, such that

$$x_1 < X_1 < x_2 \quad \text{and} \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(X_1),$$

while

$$x_2 < X_2 < x_3 \quad \text{and} \quad \frac{f(x_3) - f(x_2)}{x_3 - x_2} = f'(X_2).$$

But $X_1 < X_2$ implies that $f'(X_1) < f'(X_2)$, by our assumption that $f'(x)$ increases as $x$ increases. Therefore

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} < \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

If we interpret this inequality in terms of slopes, using notation as in Fig. 3-2, we see that it means

$$\text{slope of } P_1P_2 < \text{slope of } P_2P_3.$$

This means, however, that $P_2$ is below the line joining $P_1$ and $P_3$, so that the curve is indeed concave upward, as in Fig. 3-2.

Now, we know that when the derivative of a quantity is positive, the quantity increases as the independent variable increases. The derivative of $f'(x)$ is $f''(x)$. *Hence, if* $f''(x) > 0$ *on an interval, the first derivative is increasing and the curve is concave upward.*

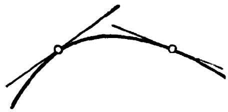


Fig. 3-4

For concavity downward, everything is reversed. The slope of the tangent decreases (i.e., the tangent turns *clockwise*) as the point of tangency moves to the right (see Fig. 3-4). This will occur if $f''(x) < 0$.

We summarize the discussion in a theorem.

THEOREM 3-F. *On an interval where* $f''(x) > 0$ *the graph of* $y = f(x)$ *is concave upward. On an interval where* $f''(x) < 0$ *the graph is concave downward.*

*Points of Inflection*

A point of inflection of a curve is a point at which the sense of concavity changes, the curve being concave upward in some interval extending from the point on one side, and the curve being concave downward in an interval extending from the point on the other side (see Fig. 3-5). The tangent is below the curve when it is concave upward, and above the curve when it is concave downward. The tangent at a point of inflection crosses the curve.

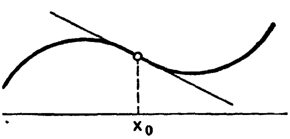



Fig. 3-5

If, as we go from left to right, the concavity is first upward and then downward, the slope at the point of inflection is a maximum. On the other hand, if the concavity changes from downward to upward (as in Fig. 3-5), the slope at the point of inflection is a minimum. In either case, since the slope has an extreme value, the derivative of the slope must be zero, by Theorem 2-B, § 2-1. That is, if $x_0$ is the point of inflection, we must have $f''(x_0) = 0$. We thus have the rule: *In dealing with a function which has a second derivative at all points under consideration, points of inflection (if any) will be among the points found by solving the equation* $f''(x) = 0$. The rule does *not* assert that every $x$ for which $f''(x) = 0$ furnishes a point of inflection. This is demonstrated by looking at the graph of $y = x^4$. Here $y'' = 12x^2$. The curve is concave upward if $x > 0$ and also if $x < 0$, so there is no point of inflection at $x = 0$. And yet $y'' = 0$ when $x = 0$.

**Example 2:** Draw the graph of $y = (x-1)^4(x-6)$ and locate all points of inflection.

We compute

$$y' = (x-1)^4 \cdot 1 + (x-6)\cdot 4(x-1)^3 = (x-1)^3[x-1+4(x-6)]$$
$$= (x-1)^3[5x-25] = 5(x-1)^3(x-5).$$
$$y'' = 5(x-1)^3 \cdot 1 + 5(x-5)\cdot 3(x-1)^2$$
$$= 5(x-1)^2[x-1+3(x-5)] = 5(x-1)^2[4x-16]$$
$$= 20(x-1)^2(x-4).$$

From the expression for $y''$ we see that it is negative when $x < 1$, negative when $1 < x < 4$, and positive when $4 < x$. It is 0 when $x = 1$ or $x = 4$. The curve is concave downward when $x < 1$ and when $1 < x < 4$; it is concave upward when $4 < x$. Hence the only point of inflection is at $x = 4$. From the expression for $y'$ we see that, as $x$ increases, $y$ increases when $x < 1$, decreases when $1 < x < 5$, and increases when $5 < x$. With a small table of values and the information gained from $y'$ and $y''$ we plot the graph, using a modified scale on the $y$-axis (see Fig. 3-6).

| $x$ | $y$ |
|---|---|
| −1 | −112 |
| 0 | −6 |
| 1 | 0 |
| 2 | −4 |
| 4 | −162 |
| 5 | −256 |
| 6 | 0 |

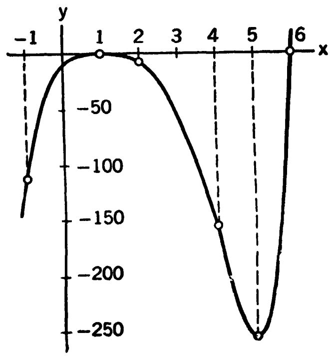



Fig. 3-6

## EXERCISES

1. In each part of this exercise an expression is given for the second derivative of a function. From this expression locate the intervals in which the graph is concave upward and those in which it is concave downward. Which values of $x$ correspond to points of inflection?
   (a) $y'' = 12x^2 - 36x$.
   (b) $y'' = 20x^3 + 6x$.
   (c) $y'' = 32x^4 - 8x^2$.
   (d) $y'' = (x^2 - 1)(x + 1)$.
   (e) $y'' = x(x-1)^2(x-2)^3$.
   (f) $y'' = (x-1)^2(x+3)(x^2-4)^2(x^3+8)$.

2. Draw the graph of each curve and locate all points of inflection.
   (a) $y = x^3 - 6x^2 + 9x + 1$.
   (b) $y = x^4 - 8x^3 + 64x + 8$.
   (c) $y = 2x^6 + 3x^5 + 10x$.
   (d) $y = 3x^4 - 2x^3 - 12x^2$.
   (e) $y = 3x^4 - 12x^3 + 12x^2 - 4$.
   (f) $y = 40x^6 + 16x^5 - x^4$.

3. Find the points of inflection of each curve. Indicate the sense of the concavity on the left and right of each point of inflection.

(a) $y = \dfrac{1}{x^2 + a^2}$.

(b) $y = \dfrac{x}{x^2 + a^2}$.

(c) $y = \dfrac{x}{1 - x^2}$.

(d) $y = \dfrac{80}{3x^4 + 80}$.

(e) $y = \dfrac{x^3}{(x - 2)^2}$.

4. Find a general formula for $y^{(n)}$ in each case.

(a) $y = (1 - x)^{-1}$.

(b) $y = (1 + x)^{-2}$.

(c) $y = (3 - 2x)^{-2}$.

(d) $y = (ax + b)^{-1}$.

5. Make a diagram showing how the graph of $y = f(x)$ might appear if $f$ has a second derivative for each $x$, given that $f(-3) = 4, f(-1) = 1, f(0) = 2$, $f''(x) > 0$ when $x < 0$ and $f''(x) < 0$ when $x > 0$, supposing in addition (a) that $f(2) = 0$; (b) that $f'(x) > 0$ and $f(x) < 4$ when $x > 0$. Why is $f'(0) = 0$ impossible in both cases?

6. Draw a diagram illustrating the behavior of a function $f$ having a second derivative for each value of $x$ and such that $f\left(\pm\dfrac{1}{n}\right) = \dfrac{1}{n^3}$ if $n$ is an even positive integer, while $f\left(\pm\dfrac{1}{n}\right) = \dfrac{1}{n^4}$ if $n$ is an odd positive integer. What must $f(0)$ be? Is $x = 0$ a point of inflection? Can you say anything about the graph being concave upward or downward in the vicinity of $x = 0$?

## 3-5 Graphing Rational Functions

Rational functions were defined in § 1-8, just after Example 8. In considering a rational function one should try to make sure that the numerator and denominator have no roots (and hence factors) in common. Thus, instead of

$$\frac{x^2 - 4}{x^3 + 8} = \frac{(x - 2)(x + 2)}{(x + 2)(x^2 - 2x + 4)}$$

we should consider

$$\frac{x - 2}{x^2 - 2x + 4}.$$

In the latter form the function is defined when $x = -2$, whereas in the first form it is not.

A rational function is continuous for each value of $x$ for which it is defined, and it is defined except for those values of $x$ which are roots of the polynomial in the denominator. In graphing a rational function it is very important to see how the values of the function vary as $x$ comes close to a root of the denominator. We assume that common factors of numerator and denominator have been removed, so that the numerator does not have a root when the denominator does. To determine the behavior of the function, we need to consider the way in which a fraction varies when its denominator approaches zero and its numerator approaches a nonzero limit. The behavior of the fraction then depends on two things: (1) the sign of the nonzero limit in the numerator, and (2) the sign of the denominator as it approaches zero. If these two signs are the same, the fraction becomes very large and positive, while if the signs are opposite, the fraction becomes very large and negative (i.e., large in absolute value, negative in sign).

**Example 1:** As $x \to 0$, $\dfrac{x-7}{x^2}$ becomes very large and negative.

**Example 2:** As $x \to 4$, $\dfrac{x-1}{(x-4)^2}$ becomes very large and positive.

As far as the sign of the denominator is concerned, this *may* depend on the direction from which $x$ approaches the root in question. We have one situation if the root is of odd order, and another if the root is of even order. Examples 1 and 2 illustrate the case for roots of even order.

**Example 3:** Consider $f(x) = \dfrac{x^2 - 2x}{(x-5)(x+1)}$ as $x \to 5$.

When $x$ is very near 5, the numerator is near $25 - 10 = 15$, and so is certainly positive. The denominator is near $6(x-5)$, and this is positive or negative according as $x > 5$ or $x < 5$. The value of $f(x)$ is near $15/6(x-5)$. Hence, if $x \to 5$ and $x$ remains greater than 5, $f(x)$ becomes very large and positive, while if $x \to 5$ and $x$ remains less than 5, $f(x)$ becomes very large and negative. The phrase "becomes very large" is not wholly adequate as a description of what occurs. Actually, $f(x)$ can be made as large as we please by making $x$ sufficiently close to 5, and the sign of $f(x)$ is controlled by the sign of $x - 5$. The implication for the part of the graph of $y = f(x)$ near $x = 5$ is indicated in Fig. 3-7. This graph also shows the general appearance of the graph near $x = -1$. When $x$ is near $-1$ we see that $f(x)$ is near $\dfrac{1-2(-1)}{-6(x+1)} = \dfrac{1}{-2(x+1)}$, so that $f(x)$ is large and negative if $x$ is near $-1$ and $x > -1$, while $f(x)$ is large and positive if $x$ is near $-1$ and $x < -1$.

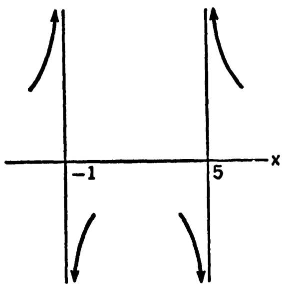


**Fig. 3-7**

A discussion such as that in Example 3 is clarified and systematized by the introduction of terminology and symbolism concerning "infinity" and "becoming infinite." If $f(x)$ is positive, and becomes larger and larger as $x \to x_0$ (no matter from which side $x$ approaches $x_0$), we say that $f(x)$ *becomes positively infinite,* or that $f(x)$ *approaches plus infinity,* as $x$ approaches $x_0$. In symbols we write $f(x) \to +\infty$ as $x \to x_0$, or

$$\lim_{x \to x_0} f(x) = +\infty.$$

The precise understanding here is that we can bring and maintain the value of $f(x)$ above (larger than) any preassigned positive number by making the absolute value $|x - x_0|$ sufficiently small (but not 0). For instance, to insure that $4/(x - 2)^2 > 10{,}000$, it suffices to have $4/10{,}000 > (x - 2)^2 > 0$, or $0 < |x - 2| < 1/50$.

In situations where $x$ approaches $x_0$ from one side only, we indicate this by writing $x \to x_0^+$ when $x - x_0$ stays positive, and $x \to x_0^-$ if $x - x_0$ stays negative (approach of $x$ to $x_0$ from right and left, respectively). The meaning of $f(x) \to +\infty$ as $x \to x_0^+$ or $x \to x_0^-$ is explained much as in the case of the unrestricted $x \to x_0$.

When $f(x)$ becomes larger and larger with negative sign, this is indicated by writing $f(x) \to -\infty$. In this case we say that $f(x)$ "approaches negative infinity" or "becomes negatively infinite." This may occur as $x \to x_0$ or as $x \to x_0^+$ or $x \to x_0^-$. For instance, when $f(x) = \dfrac{x - 7}{x^2}$, $f(x) \to -\infty$ as $x \to 0$. When $f(x) = \dfrac{x^2 - 2x}{(x - 5)(x + 1)}$, $f(x) \to +\infty$ as $x \to 5^+$, while $f(x) \to -\infty$ as $x \to 5^-$.

The symbols $+\infty$ and $-\infty$ are also used in connection with the independent variable, in describing how $f(x)$ behaves as $x$ becomes very large and positive ($x \to +\infty$) or very large and negative ($x \to -\infty$).

**Example 4:** If $f(x) = \dfrac{4x}{x^2 + 1} = \dfrac{4}{x + (1/x)}$, $f(x) \to 0^+$ as $x \to +\infty$, and $f(x) \to 0^-$ as $x \to -\infty$. That is, $f(x)$ is positive and very small when $x$ is large and positive, and $f(x)$ is negative and very small when $x$ is large and negative.

We wish to emphasize that the symbols $+\infty$ and $-\infty$ have been used to describe certain things in connection with the behavior of variables. The symbols $+\infty$, $-\infty$ are *not* real numbers. We are not attempting to treat them as numbers. That is, we do not attempt to do addition, subtraction, multiplication, or division with these symbols. Our only use of the symbols $+\infty$, $-\infty$ and of the notion of "approaching infinity" is in connection with statements about limits.

Now suppose we are constructing the graph of the fractional rational function

$$y = f(x) = \frac{p(x)}{P(x)},$$

where $p(x)$ and $P(x)$ are polynomials without common roots. If $x_0$ is a real root of $P(x)$, the line $x = x_0$ is called an *asymptote* of the graph. (Later on we shall explain more fully the general relationship between a curve and an asymptote of the curve.) As $x$ approaches $x_0$ from one side, $f(x)$ approaches either $+\infty$ or $-\infty$, and the corresponding part of the graph flattens out toward one end of the asymptote. If the root is of odd order, the curve approaches opposite ends of the asymptote as $x$ approaches $x_0$ from the two different sides. But if the root is of even order, the curve approaches the same end of the asymptote from the two different sides. This is illustrated most simply by

$$y = \frac{1}{x} \quad \text{(root of order 1 at } x = 0\text{)}$$

and

$$y = \frac{1}{x^2} \quad \text{(root of order 2 at } x = 0\text{)}.$$

See Fig. 3-8.

An asymptote $x = x_0$ (parallel to the $y$-axis) is called a *vertical* asymptote. The graph of a fractional rational function may also have a *horizontal* asymptote (i.e., one parallel to the $x$-axis). In fact, there is such an asymptote (and only one) if and only if the degree of the numerator *does not exceed* the degree of the denominator. We discover this kind of asymptote by considering what happens as $x \to +\infty$ or $x \to -\infty$. Each of the two graphs in Fig. 3-8 has the line $y = 0$ as a horizontal asymptote.

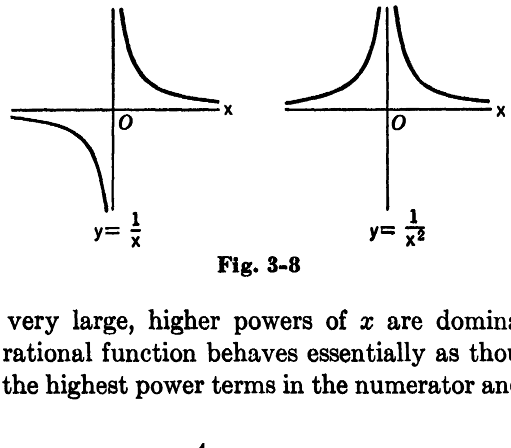



**Fig. 3-8**

When $x$ is very large, higher powers of $x$ are dominant over lower powers, and a rational function behaves essentially as though we were to discard all but the highest power terms in the numerator and denominator, respectively.

**Example 5:** Let $f(x) = \dfrac{4x}{x^2 + 1}$. For large values of $x$ we have, approximately,

$$f(x) \sim \frac{4x}{x^2} = \frac{4}{x}.$$

Hence $f(x) \to 0^+$ as $x \to +\infty$, and $f(x) \to 0^-$ as $x \to -\infty$. The line $y = 0$ is a horizontal asymptote, and the behavior of the graph near the asymptote is indicated in Fig. 3-9.

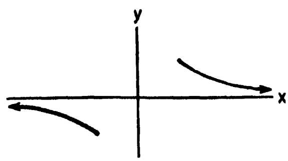



**Fig. 3-9**

**Example 6:** Let $f(x) = \dfrac{2x^2 + x - 3}{4x - x^2}$. For large values of $x$ we have, approximately,

$$f(x) \sim \frac{2x^2}{-x^2} = -2.$$

Thus $f(x) \to -2$ as $x \to +\infty$ and also as $x \to -\infty$. The line $y = -2$ is a horizontal asymptote.

If the degree of the numerator *exceeds* that of the denominator, the behavior of the rational function for large values of $x$ may be determined by using long division to express the function as the sum of a polynomial and a proper rational function, i.e., one in which the degree of the numerator is less than that of the denominator.

**Example 7:** $y = \dfrac{3x^3 - x^2 - 2}{4x^2 - 8x}$.

The long division is indicated:

$$\begin{array}{r|l}
 & \tfrac{3}{4}x + \tfrac{5}{4} \\ \hline
4x^2 - 8x & 3x^3 - x^2 \qquad\qquad - 2 \\
 & 3x^3 - 6x^2 \\ \cline{2-2}
 & \qquad 5x^2 \qquad\qquad - 2 \\
 & \qquad 5x^2 - 10x \\ \cline{2-2}
 & \qquad\qquad 10x - 2
\end{array}$$

So we have
$$y = \frac{3}{4}x + \frac{5}{4} + \frac{10x - 2}{4x^2 - 8x}.$$

For large values of $x$ the fractional term is approximately $10x/4x^2 = 5/2x$. Thus approximately,

$$y \sim \frac{3}{4}x + \frac{5}{4} + \frac{5}{2x}.$$

This indicates that the graph is close to the straight line

$$y = \frac{3}{4}x + \frac{5}{4};$$

it is above the line if $x$ is large and positive, and below the line if $x$ is large and negative, for the deviation from the line is approximately $5/2x$. The line $y = \frac{3}{4}x + \frac{5}{4}$ is an oblique asymptote of the graph. The relation of the graph to the asymptote is indicated in Fig. 3-10. We do not show what happens for intermediate values of $x$.

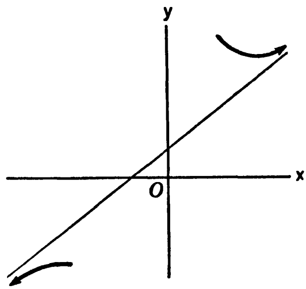



**Fig. 3-10**

We shall now summarize the suggestions as to procedure in drawing the graph of a fractional rational function:

(a) Locate all the vertical asymptotes and indicate how the graph looks near these asymptotes.

(b) Examine the behavior of the function as $x \to +\infty$ and $x \to -\infty$. This examination *may* disclose a horizontal asymptote or an oblique asymptote. If the degree of the numerator exceeds that of the denominator by two or more, there will be no horizontal nor oblique asymptote.

(c) Sketch in the rest of the graph, using any conveniently available information such as: where the numerator is zero, where the derivative of the function is zero, and where the derivative is positive or negative.

We point out specifically that the graph *may* intersect a horizontal or an oblique asymptote. The plane is divided into several compartments by the vertical asymptotes; the part of the graph in any one of these compartments is one continuous piece, but it does not connect with the piece in a different compartment.

It is not always necessary to use the derivative to get a pretty fair notion of the appearance of the graph. The use of the second derivative is much less essential than that of the first derivative, and in many cases it is not worth the trouble of computing it.

There is one special matter that deserves comment: *the inspection for symmetry*. There are two types of symmetry that are easily detected.

(1) *Symmetry with respect to the y-axis*. The graph has this kind of symmetry if all the occurring powers of $x$ are even, so that $f(x)$ is the same as $f(-x)$. This means that if we fold the plane along the $y$-axis, the part of the graph for which $x > 0$ will fall exactly on top of the part for which $x < 0$.

(2) *Symmetry with respect to the origin*. The graph has this kind of symmetry if $f(-x) = -f(x)$. This means that, for each point on the graph, the point directly through the origin from it and an equal distance on the other side of the origin, is also on the graph. This kind of symmetry occurs if all the powers of $x$ in the numerator are odd and all those in the denominator are even, or vice versa, e.g.,

$$y = \frac{x}{x^2 + 1} \quad \text{and} \quad y = \frac{x^4 - 16}{x^3 - 2x}.$$

With either of these kinds of symmetry, we can draw the graph for $x < 0$ as soon as we have drawn it for $x > 0$.

We conclude this section by completing the graph of $y = \dfrac{x^2 - 2x}{x^2 - 4x - 5}$, the vertical asymptotes of which were indicated in Fig. 3-7. For an accurate notion of the relation of the graph to its horizontal asymptote $y = 1$ we use the alternative formula

$$y = 1 + \frac{2x + 5}{x^2 - 4x - 5},$$

which is obtained by long division. This indicates that $y > 1$ if $x$ is large and positive, while $y < 1$ if $x$ is large and negative. Hence the curve is above the asymptote $y = 1$ at the extreme right, and below it at the extreme left. We note also that $y = 0$ if $x = 0$ and if $x = 2$. When these facts are combined with what we know from Fig. 3-7, we are able to draw

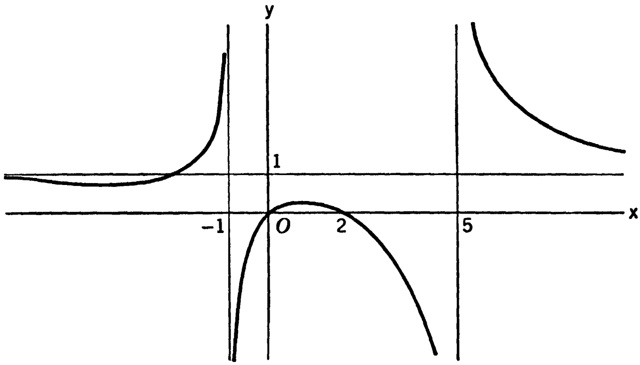



**Fig. 3-11**

the graph much as in Fig. 3-11. Observe that the graph must cross the asymptote $y = 1$ somewhere to the left of $x = -1$. The crossing is at $x = -\frac{5}{2}$. We expect to find two points where the tangent is horizontal: one somewhere to the left of $x = -1$, and one between $x = 0$ and $x = 2$. We find these points exactly by using the derivative. This derivative is

$$y' = \frac{-2(x^2 + 5x - 5)}{(x^2 - 4x - 5)^2};$$

verification is left to the student. The points of zero slope occur when $x^2 + 5x - 5 = 0$, or

$$x = \frac{-5 \pm 3\sqrt{5}}{2}.$$

The approximate values are $x = -5.85, 0.85$. In the graph the scale is slightly distorted for $x$ near $-5.85$, in order to show the trend of the curve. Actually, the curve is very close to the asymptote for all values of $x$ less than $-\frac{5}{2}$.

We now say a bit more about asymptotes. A straight line is called an asymptote of a curve if, as a point moves out along an extremity of the curve, its distance from the line approaches zero, and if the tangent to the curve at the point approaches coincidence with the straight line. Thus the extremity of the curve becomes more and more indistinguishable, both in position and direction, from the extremity of the line as we move out along the curve. The curve may cross the asymptote.

## EXERCISES

**1.** Graph each function. Locate all asymptotes, and note symmetry, if any. Use the first derivative.

(a) $y = \dfrac{x-1}{(x-2)^2}$.

(b) $y = \dfrac{x-2}{(x+1)^2}$.

(c) $y = \dfrac{x^2}{x^2-4}$.

(d) $y = \dfrac{x^2+x}{(x-2)^2}$.

(e) $y = \dfrac{x^2-4}{(x-1)^2}$.

(f) $y = \dfrac{4x}{x^2-9}$.

(g) $y = \dfrac{x-4}{x(x-3)}$.

(h) $y = \dfrac{(x-3)^2}{2+x-x^2}$.

(i) $y = \dfrac{x^2-4}{x^3}$.

(j) $y = \dfrac{x^2}{(x-2)^3}$.

**2.** Follow the instructions of Exercise 1.

(a) $y = \dfrac{(x-2)^2}{x-1}$.

(b) $y = \dfrac{x^2-x-2}{x-1}$.

(c) $y = \dfrac{x(x+2)}{x-2}$.

(d) $y = \dfrac{x^2+16}{x-3}$.

(e) $y = \dfrac{6(x-4)}{3x^2+2x-8}$.

(f) $y = \dfrac{(x-2)^3}{x+2}$.

(g) $y = \dfrac{x^3-4x^2}{(x-2)^2}$.

(h) $y = \dfrac{(x+1)^3}{x^2-2x}$.

**3.** Graph each function as well as you can without using the derivative. Can you tell how many horizontal tangents there are without actually finding them?

(a) $y = \dfrac{x^2}{(x+1)^2(x-2)}$.

(b) $y = \dfrac{x}{(x+1)(x-2)^2}$.

(c) $y = \dfrac{(x+1)^2}{x^2(x-3)}$.

(d) $y = \dfrac{x+1}{x(x-1)(x-2)}$.

(e) $y = \dfrac{x+1}{x(x^2-4)}$.

(f) $y = \dfrac{x^3-8}{x(x^2-9)}$.

(g) $y = \dfrac{6(x+4)}{x(x-2)^2(x+3)}$.

(h) $y = \dfrac{3(x-1)^3}{2(x^2-x-2)}$.

(i) $y = \dfrac{3x^3-x^2-2}{4x^2-8x}$.

(j) $y = \dfrac{(x+1)(x^2-1)}{x(x-2)}$.

(k) $y = \dfrac{(x+2)^2(x+1)}{x(x-2)}$.

(l) $y = \dfrac{x^2(x-1)}{(x+1)(x-3)}$.

## 3-6 Fractional Exponents

Thus far we have not employed fractional exponents in this book. We have assumed that the student is acquainted with the use of positive and negative integers and zero as exponents, according to the rules

$$a^{-n} = \frac{1}{a^n}, \quad a^0 = 1 \qquad (a \neq 0).$$

Students are also expected to have some familiarity with fractional exponents, and to be able to use the exponent laws

$$a^m a^n = a^{m+n},$$

$$(a^m)^n = a^{mn},$$

$$(ab)^n = a^n b^n$$

for fractional as well as integral exponents. Our main purpose in this section is to discuss the function $y = x^n$ when $n$ is a fractional exponent, and to prove the validity of the exponent rule of differentiation in this case:

$$\frac{d}{dx} x^n = nx^{n-1}.$$

The meaning of a fractional exponent is defined in terms of the *root* concept, so we begin with a discussion of $q$th roots, where $q$ is a positive integer. Let $x$ and $y$ be real numbers. We call $y$ a $q$th root of $x$ if $y^q = x$. Evidently 0 is the only $q$th root of 0. If $x < 0$ and $q$ is even, there is no real $q$th root of $x$, for an even power of a real number cannot be negative. Two questions arise: (1) For a given $q$, which numbers $x$ do possess $q$th roots? (2) How many $q$th roots are there for a given $x$? In this discussion we consider real numbers only. We shall show that every real $x$ has exactly one $q$th root if $q$ is odd, and that if $q$ is even, every positive $x$ has exactly two $q$th roots, one of them being the negative of the other. To show these things we study the graph of the equation $x = y^q$, regarding $y$ as independent and $x$ as dependent (reversing the usual roles of these letters).

For $x = y^q$ we have $dx/dy = qy^{q-1}$. We consider odd and even $q$ separately. We assume $q > 1$, for $x = y$ when $q = 1$, and everything is clear in this case.

*q odd:* In this case $q - 1$ is even, and $dx/dy$ is never negative. Hence $x$ increases steadily as $y$ increases, with $x \to -\infty$ as $y \to -\infty$ and $x \to +\infty$ as $y \to +\infty$. The graph appears in Fig. 3-12. Note the labeling of the axes.

*q even:* In this case $q - 1$ is odd, and the sign of $dx/dy$ is the same as that of $y$. The graph has the appearance shown in Fig. 3-13. It is symmetric with respect to the $x$-axis.

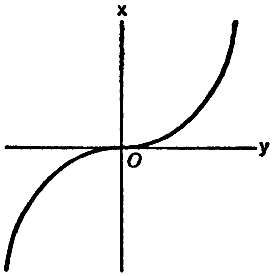



**Fig. 3-12**

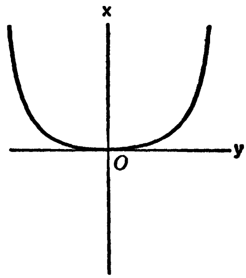



**Fig. 3-13**

Now consider the question of $q$th roots. From Fig. 3-12 we see that when $q$ is odd and $x$ is any given number, there is exactly one $y$ such that $y^q = x$. This $y$ is determined by $x$, and so we may regard $y$ as a function of $x$, instead of regarding $x$ as a function of $y$. This $q$th root of $x$ is denoted by $y = \sqrt[q]{x}$. The meaning of the fractional exponent $1/q$ is then given by the definition

$$x^{1/q} = \sqrt[q]{x}.$$

To graph $y = x^{1/q}$, with $y$ dependent and $x$ independent, we have only to reorient the graph in Fig. 3-12, allowing for the change in the position of the axes. It is as though Fig. 3-12 were drawn on transparent paper and then viewed from the other side of the paper (see Fig. 3-14). When $q$ is

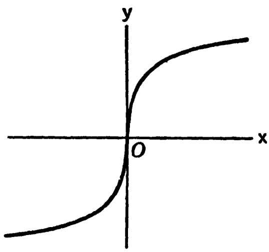



**Fig. 3-14**

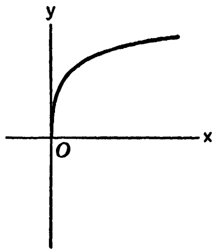



**Fig. 3-15**

even, the situation is a bit different. If $x > 0$ there are two values of $y$ such that $y^q = x$. We choose the positive value of $y$ and call it the *principal* $q$th root of $x$; we denote it by $y = \sqrt[q]{x}$, and again we write $x^{1/q} = \sqrt[q]{x}$. The graph of $y = x^{1/q}$ when $q$ is even is shown in Fig. 3-15. The curve stops at the origin.

Other fractional powers of $x$ are defined as follows:

$$x^{p/q} = (x^{1/q})^p = (\sqrt[q]{x})^p,$$

where $q$ is a positive integer and $p$ is any positive or negative integer. This definition applies to any positive $x$, and to any negative $x$ if $q$ is odd. It can be shown that the usual laws of exponents apply to fractional as well as to integral exponents.

Now we turn to the question of differentiating fractional powers of $x$. First we observe that the function $x^{1/q}$ is certainly differentiable if $x \neq 0$ (and if $x > 0$ when $q$ is even). For, differentiability of $y = x^{1/q}$ with respect to $x$ means the same as the graph having a tangent not parallel to the $y$-axis. The graph of $y = x^{1/q}$ is the same (except for orientation) as the graph of $y^q = x$, and the latter graph certainly has a tangent not parallel to the $y$-axis if $y \neq 0$, because $y^q$ is differentiable (with respect to $y$) with a nonzero derivative. Since $x^{1/q}$ is differentiable, so is its $p$th power, which is $x^{p/q}$. For actual computation of the derivative, let

$$y = x^{p/q}, \quad \text{and hence} \quad y^q = x^p,$$

or

$$(x^{p/q})^q = x^p.$$

Then, using the composite function rule and the exponent rule of differentiation for integers, we have

$$q(x^{p/q})^{q-1} \frac{d}{dx}(x^{p/q}) = px^{p-1},$$

or

$$\frac{d}{dx}(x^{p/q}) = \frac{px^{p-1}}{q(x^{p/q})^{q-1}} = \frac{px^{p-1}}{qx^{p-(p/q)}} = \frac{p}{q}\, x^{(p/q)-1}.$$

Hence the exponent rule of differentiation also applies to fractional powers of $x$.

**Example 1:** $\frac{d}{dx} x^{1/2} = \frac{1}{2}x^{-1/2}$. This is often used in the form

$$\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}}.$$

We often deal with fractional powers of a function of $x$, rather than of $x$ itself. We then use the chain rule.

**Example 2:**

$$\frac{d}{dx}(a^2 - x^2)^{-3/2} = -\frac{3}{2}(a^2 - x^2)^{-5/2}\frac{d}{dx}(a^2 - x^2)$$
$$= -\frac{3}{2}(a^2 - x^2)^{-5/2}(-2x) = \frac{3x}{(a^2 - x^2)^{5/2}}.$$

## EXERCISES

**1.** Find $y'$ in each case.

(a) $y = 4(1 - x)^{4/3} + 2(3x^2 - 2x + 1)^{1/3} - (2x - x^2)^{-1/3}$.

(b) $y = (2x - 1)^{5/2} - (2x^3 - x^2 + 2x - 1)^{-1/2}$.

(c) $y = \sqrt{\left(\dfrac{1-x}{1+x}\right)^3}$.

(d) $y = \left(\dfrac{x^2+16}{x}\right)^{2/3}$.

(e) $y = (4 - x^{2/3})^{3/2}$.

(f) $y = x(x-2)^{3/4}(x+2)^{1/4}$.

2. Find $\dfrac{dy}{dx}$ and simplify your answer.

(a) $y = (3x - 4)(2 + 3x)^{1/2}$.

(b) $y = (15x - 2)(5x + 1)^{3/2}$.

(c) $y = x(25 - x^2)^{-1/2}$.

(d) $y = \dfrac{\sqrt{16 - x^2}}{x}$.

(e) $y = \dfrac{x - 3}{\sqrt{6x - x^2}}$.

(f) $y = \dfrac{3x + 8}{\sqrt{4 + 3x - x^2}}$.

(g) $y = (\frac{1}{5}x^2 - \frac{2}{15}a^2)(a^2 - x^2)^{3/2}$.

(h) $y = (75x^2 - 80x + 128)(4 + 5x)^{1/2}$.

(i) $y = (135x^2 - 144x + 128)(4 + 3x)^{3/2}$.

(j) $y = \dfrac{16x^3 - 24x^2 - 42x + 25}{(2 + x - x^2)^{5/2}}$.

3. In the formula $T = 2\pi g^{-1/2}(l^2 - r^2)^{1/4}$, $l$ is the length of a conical pendulum, $r$ is the radius of the path described by the bob, and $T$ is the period. (a) Find $dT/dl$ if $l = 13$ and $r = 5$. Take $g = 32$. (b) Discuss the sense of concavity of the graph of $T$ as a function of $l$.

4. In a certain electrostatic field the electric intensity $E$ at a point of the $x$-axis is $E = (x^4 + 2a^4 + a^8x^{-4})^{-1/4}$, where $a > 0$. Find $dE/dx$ and determine for which positive values of $x$ the derivative is positive.

## 3-7 Implicit Functions

Sometimes $y$ is defined as a function of $x$, not by giving the value of $y$ explicitly in terms of $x$, but by giving an equation in $x$ and $y$. Such an equation may not determine $y$ *uniquely* in terms of $x$, but in the situations we commonly meet the equation determines one or more distinct functions.

**Example 1:** The equation $x^2 + y^2 - 16 = 0$ determines two functions of $x$:

$$y = \sqrt{16 - x^2} \quad \text{and} \quad y = -\sqrt{16 - x^2}.$$

**Example 2:** The equation $5x^2 - 6xy + 5y^2 = 128$ determines two functions of $x$, which we find by solving for $y$ by the quadratic formula:

$$y = \frac{6x \pm \sqrt{36x^2 - 4(5)(5x^2 - 128)}}{10},$$

or $$y = \frac{1}{5}[3x + 4\sqrt{40 - x^2}] \quad \text{and} \quad y = \frac{1}{5}[3x - 4\sqrt{40 - x^2}].$$

**Example 3:** The equation $x^3 - 3axy + y^3 = 0 (a > 0)$ has as its graph the curve shown in Fig. 3-16. This curve is called the *folium of Descartes.* We shall

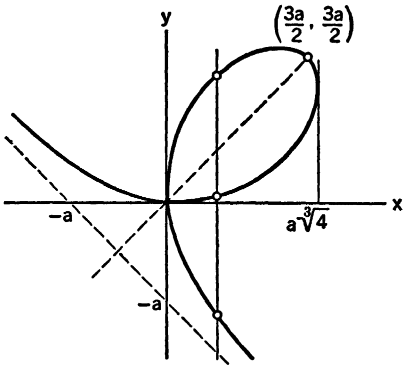



**Fig. 3-16**

not at this time go into the details of how the graph is constructed. From the graph it is evident that to each $x$ such that $0 < x < a\sqrt[3]{4}$ correspond three distinct values of $y$ such that $(x, y)$ is a point on the graph. Thus the equation determines three functions of $x$ on the interval $0 < x < a\sqrt[3]{4}$. On the other hand, if $x > a\sqrt[3]{4}$ or $x < 0$, there is just one value of $y$ for each $x$. The line $x + y = -a$ is an asymptote.

When an equation in $x$ and $y$ determines $y$ as a function of $x$ (or as one of several functions of $x$), but when we do not have a direct explicit formula of $y$ in terms of $x$, we say that $y$ is defined *implicitly* as a function of $x$ by the equation.

When $y$ is a differentiable function of $x$ which is defined implicitly by an equation of a suitable type, we can calculate the derivative $dy/dx$ directly from the equation without solving for $y$ explicitly. The derivative will usually be expressed in terms of both $x$ and $y$ rather than in terms of $x$ alone. In finding $dy/dx$ we use the composite function rule and regard $y$ as a function of $x$ wherever it appears. For instance,

$$\frac{d}{dx}(y^2) = 2y\frac{dy}{dx},$$

$$\frac{d}{dx}(x^2y^3) = x^2 \cdot 3y^2\frac{dy}{dx} + 2xy^3.$$

**Example 4:** From $x^2 + y^2 - 16 = 0$ we have

$$2x + 2y\frac{dy}{dx} = 0, \quad \text{or} \quad \frac{dy}{dx} = -\frac{x}{y}.$$

In this case we can solve for $y$ explicitly and check our result:

$$y = \pm\sqrt{16 - x^2} = \pm(16 - x^2)^{1/2},$$

$$\frac{dy}{dx} = \pm\frac{1}{2}(16 - x^2)^{-1/2}(-2x) = \frac{-x}{\pm\sqrt{16 - x^2}} = -\frac{x}{y}.$$

**Example 5:** From the equation $x^3 - 3axy + y^3 = 0$ in Example 3 we have

$$3x^2 - 3ax\frac{dy}{dx} - 3ay + 3y^2\frac{dy}{dx} = 0,$$

$$\frac{dy}{dx} = \frac{ay - x^2}{y^2 - ax}.$$

If we want a numerical result we must put in the coordinates $(x, y)$ of a point on the graph. For instance, if $x = \frac{2}{3}a$, the three corresponding values of $y$ are $\frac{4}{3}a$, $\frac{a}{3}(\sqrt{6} - 2)$, and $\frac{a}{3}(-\sqrt{6} - 2)$. At the point $\left(\frac{2}{3}a, \frac{4}{3}a\right)$ the slope of the curve is

$$\frac{dy}{dx} = \frac{\frac{4a^2}{3} - \frac{4a^2}{9}}{\frac{16a^2}{9} - \frac{2a^2}{3}} = \frac{4}{5}.$$

We can also compute second derivatives of functions which are defined implicitly.

**Example 6:** Find $\frac{d^2y}{dx^2}$ from $3x^2 + 4y^2 = 12$, without solving for $y$.

First we have

$$6x + 8y\frac{dy}{dx} = 0, \quad \text{or} \quad \frac{dy}{dx} = -\frac{3x}{4y}.$$

Then
$$\frac{d^2y}{dx^2} = -\frac{3}{4}\,\frac{y\cdot 1 - xy'}{y^2}.$$

But
$$y' = \frac{dy}{dx} = -\frac{3x}{4y}, \quad \text{so} \quad xy' = -\frac{3x^2}{4y},$$

and
$$y'' = \frac{d^2y}{dx^2} = -\frac{3}{4}\,\frac{y + \frac{3x^2}{4y}}{y^2} = -\frac{3}{16}\,\frac{4y^2 + 3x^2}{y^3}.$$

This result may be simplified, since $4y^2 + 3x^2 = 12$. Thus

$$y'' = -\frac{3}{16}\cdot\frac{12}{y^3} = -\frac{9}{4y^3}.$$

## EXERCISES

**1.** Differentiate each of the following expressions with respect to $x$, treating $y$ as a function of $x$ not explicitly known.

(a) $xy^3$; (b) $\sqrt{xy}$; (c) $\frac{y^2}{x^2}$; (d) $x^{1/2}y^{-3/2} + x^{3/2}y^{-1/2}$.

**2.** Find the slope of each curve at the point or points indicated. Also find $\frac{d^2y}{dx^2}$ at these points.

(a) $25x^2 - 16y^2 = 400$ at $(5, \frac{15}{4})$ and $(5, -\frac{15}{4})$.
(b) $9x^2 + 25y^2 = 225$ at $(-3, \frac{12}{5})$ and $(3, \frac{12}{5})$.
(c) $2x^2 - xy + 3y^2 = 18$ at $(3, 1)$.
(d) $5x^2 - 6xy + 5y^2 = 128$ at $x = y = 4\sqrt{2}$.
(e) $x^2 + xy + y^2 = 4$ at $(2, -2)$.
(f) $x^3 + y^3 = 18xy$ at $(8, 4)$.

**3.** Find $\frac{dy}{dx}$ in terms of $x$ and $y$ from each equation.

(a) $y^3 = 4(x^2 + y^2)$.
(b) $y^4 + x^2y^2 = 100$.
(c) $x^{2/3} + y^{2/3} = a^{2/3}$.
(d) $x^{1/2} + y^{1/2} = a^{1/2}$.

**4.** Find $y'$ and $y''$ without explicitly solving for $y$.

(a) $25x^2 - 16y^2 = 400$.
(b) $x^3 + y^3 = a^3$.
(c) $x^4 + y^4 = a^4$.
(d) $x^{1/2} + y^{1/2} = 1$.
(e) $x^3 + y^3 - 3axy = 0$.

**5.** Prove that the curves $x^2 + 3y^2 = 24$ and $3x^2 - y^2 = 12$ intersect at right angles at the point $(\sqrt{6}, \sqrt{6})$.

**6.** The total surface area of a right circular cone of height $h$ and radius of base $r$ is $S = \pi(r^2 + r\sqrt{r^2 + h^2})$. If $S$ is constant, find $\frac{dr}{dh}$ when $r = 3$ and $h = 4$.

## 3-8 Circles and Ellipses

*Circles*

In the first part of § 1-5 we saw that the equation

$$(x - a)^2 + (y - b)^2 = r^2 \tag{1}$$

describes the circle of radius $r$ with center at $(a, b)$. In particular, if the center is at the origin, the equation takes the form

$$x^2 + y^2 = r^2.$$

If we write the equation (1) in the expanded form

$$x^2 - 2ax + a^2 + y^2 - 2by + b^2 = r^2,$$

we see that it contains $x^2 + y^2$, terms of first power in $x$ and $y$, and constant terms. That is, the equation is of the general type

$$x^2 + y^2 + 2Ax + 2By + C = 0, \tag{2}$$

where $A$, $B$, and $C$ are constants. If we have the equation given to us in this latter form, we can locate the center of the circle and find its radius by a process of completing the squares.

**Example 1:** Consider the equation

$$x^2 + y^2 - 6x + 4y - 12 = 0.$$

We write it in the form

$$x^2 - 6x \qquad + y^2 + 4y \qquad = 12.$$

Then we add the terms necessary for the completion of the squares, and compensate by adding the proper amount on the right side:

$$x^2 - 6x + 9 + y^2 + 4y + 4 = 12 + 13 = 25,$$
$$(x - 3)^2 + (y + 2)^2 = 5^2.$$

The equation is that of a circle of radius 5 and center at $(3, -2)$.

If this procedure is applied to the general equation (2), we obtain

$$(x + A)^2 + (y + B)^2 = A^2 + B^2 - C. \tag{3}$$

This equation represents a circle if $A^2 + B^2 - C > 0$. If $A^2 + B^2 - C < 0$, however, the equation has no graph (that is, there are no points which satisfy the equation), for the left side of equation (3) can never be equal to a negative number. In the special case that $A^2 + B^2 - C = 0$, the graph of the equation consists of just one point:

$$x = -A, \qquad y = -B.$$

By using either equation (1) or (2) we can find a circle that fulfills certain conditions. For example, if three points do not lie on the same straight line, there is a unique circle which passes through all three points. We find this circle by solving three simultaneous linear equations in which the coefficients $A$, $B$, $C$ of (2) are unknowns.

**Example 2:** Find the circle through the points $(4, 2)$, $(1, 3)$, and $(-3, -5)$. We substitute into (2):

$$16 + 4 + 8A + 4B + C = 0, \quad \text{or} \quad 8A + 4B + C = -20,$$
$$1 + 9 + 2A + 6B + C = 0, \quad \text{or} \quad 2A + 6B + C = -10,$$
$$9 + 25 - 6A - 10B + C = 0, \quad \text{or} \quad 6A + 10B - C = 34.$$

We solve by successive elimination:

$$6A - 2B = -10 \quad \text{(from the first two equations)},$$
$$8A + 16B = 24 \quad \text{(from the last two equations)}.$$

Now multiply the first of these by 4, the second by $\frac{1}{2}$, and add:

$$28A = -28, \quad \text{or} \quad A = -1.$$

Going back, we find $B = 3A + 5 = 2$, $C = -2A - 6B - 10 = -20$. The equation of the circle is therefore

$$x^2 + y^2 - 2x + 4y - 20 = 0.$$

We leave it for the student to find the center and radius and check the results on a diagram.

The slope of the tangent to a circle at a specified point can be found by calculus, using the derivative. It can be found also as the negative reciprocal of the slope of the radius to the point of tangency.

In some problems the method of procedure reveals itself most naturally after we draw a figure and study it. For example, if it is required to find the circle which passes through two given points $P_1$, $P_2$ and has its center on a given line, a figure suggests that we find the center as the intersection of the given line and the perpendicular bisector of the segment joining $P_1$ and $P_2$. We can then compute the radius of the circle.

The following example illustrates how to find the points of intersection (if any) of two circles.

**Example 3:** Find the points of intersection of the two circles

$$x^2 + y^2 + 7x - 9y - 24 = 0,$$
$$x^2 + y^2 + 2x - 4y - 29 = 0.$$

To solve simultaneously we subtract one equation from the other. This gives us a linear equation:

$$5x - 5y + 5 = 0, \quad \text{or} \quad x - y + 1 = 0.$$

We then solve for $y$ (or $x$) in the linear equation, and substitute back into the equation of one of the circles: $y = x + 1$,

$$x^2 + (x + 1)^2 + 2x - 4(x + 1) - 29 = 0,$$
$$2x^2 - 32 = 0, \qquad x = \pm 4,$$
$$y = 4 + 1 = 5 \quad \text{or} \quad y = -4 + 1 = -3.$$

The points of intersection are (4, 5) and (−4, −3).

We remark that the linear equation represents the line through the two points of intersection (see Fig. 3-17). If this procedure is attempted in a

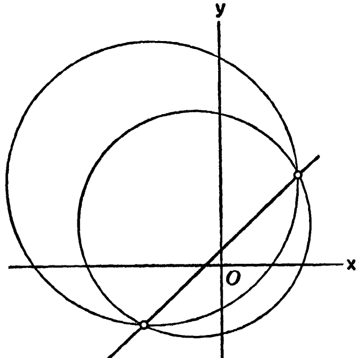



**Fig. 3-17**

case where the circles do not intersect, one arrives at a quadratic equation with no real roots.

*Ellipses*

An ellipse is a curve of great interest and importance. The appearance of an ellipse is very familiar. An oblique plane cross section of a right circular cylinder is an ellipse. When we view the rim of a drinking glass from above and to one side, it appears to have the shape of an ellipse. The planets travel around the sun in orbits which are approximately elliptical.

Our most convenient starting point is the following definition. Let $F$ and $F'$ be two distinct points. Let $2a$ be a constant larger than the distance $F'F$. Consider the curve (in a plane through the line $F'F$) composed of all points $P$ such that the sum of the distances $F'P$ and $FP$ is $2a$. This curve is called an ellipse, and each of the points $F$, $F'$ is called a *focus* of the ellipse (see Fig. 3-18).

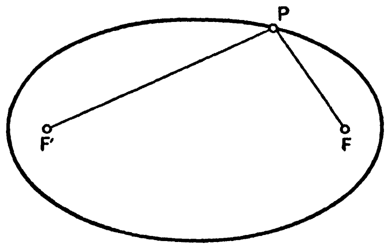



**Fig. 3-18**

We shall introduce certain standard notations for the dimensions of an ellipse. The curve is evidently symmetric with respect to the line through the foci, and also with respect to the perpendicular bisector of the line segment $F'F$. In Fig. 3-19 the segment $A'A$ is called the *major axis* and the

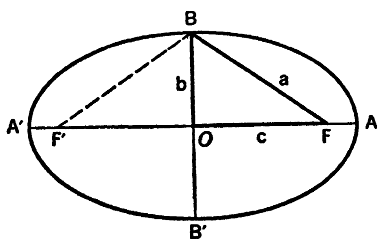



**Fig. 3-19**

segment $B'B$ is called the *minor axis* of the ellipse. We denote the length of the minor axis by $2b$ and the distance between the foci by $2c$. Evidently $BF = a$, so

$$a^2 = b^2 + c^2. \tag{4}$$

Since $F'A + FA = 2a$ and $A'F' = FA$, we see that $A'A = 2a$. That is, the length of the major axis is $2a$.

Observe that $b < a$. The ellipse can be long and thin (when $b$ is small in relation to $a$, and the foci are near the ends of the major axis), or nearly circular (when $c$ is small and $b$ is nearly as large as $a$). The ratio of $c$ to $a$ is called the *eccentricity* of the ellipse, and denoted by $e$:

$$e = \frac{c}{a} = \frac{\sqrt{a^2 - b^2}}{a}. \tag{5}$$

Observe that $0 < e < 1$. Long thin ellipses have eccentricity near 1, while nearly circular ellipses have eccentricity near 0.

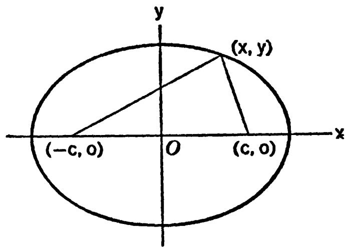



**Fig. 3-20**

The equation of an ellipse is simplest if we put the center of the curve at the origin and the major axis along either the $x$- or $y$-axis. We shall derive the equation with the foci on the $x$-axis (see Fig. 3-20). The definition of the ellipse is then expressed by the equation

$$\sqrt{(x + c)^2 + y^2} + \sqrt{(x - c)^2 + y^2} = 2a. \tag{6}$$

This equation can be made much simpler by squaring. Transpose the first radical and square both sides of the equation. After simplification we obtain

$$a\sqrt{(x + c)^2 + y^2} = a^2 + cx.$$

Now square again:

$$a^2x^2 + 2a^2cx + a^2c^2 + a^2y^2 = a^4 + 2a^2cx + c^2x^2,$$

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2).$$

In view of (4) this becomes

$$b^2x^2 + a^2y^2 = a^2b^2,$$

or

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{7}$$

We have shown that (7) is satisfied if (6) is; it can be shown, conversely, that (6) is satisfied if (7) is. We omit the details. A similar type of converse argument was given in the case of the equation of the parabola, in § 2-4. The equation (7) is therefore the equation of the ellipse in the standard position shown in Fig. 3-20.

**Example 4:** Find the equation of the ellipse with foci at $(\pm 5, 0)$ and the ends of the minor axis at $(0, \pm 12)$.

From what is given we know that $c = 5$ and $b = 12$. Hence $a^2 = 25 + 144 = 169$, $a = 13$. The equation is

$$\frac{x^2}{169} + \frac{y^2}{144} = 1.$$

The eccentricity is $e = \frac{5}{13}$.

If the ellipse has its center at the origin and its foci on the $y$-axis, the roles of $x$ and $y$ are exchanged in its equation, which is

$$\frac{x^2}{b^2} + \frac{y^2}{a^2} = 1. \tag{8}$$

Equations (7) and (8) have the same general appearance; it is the fact that $a > b$ which indicates the difference between them.

**Example 5:** The equation $25x^2 + 16y^2 = 400$ can be written in the form

$$\frac{x^2}{16} + \frac{y^2}{25} = 1.$$

It represents an ellipse with $a = 5$, $b = 4$, and foci on the $y$-axis.

It is easy to deal with an ellipse whose center is not at the origin, provided its major axis is parallel to one of the coordinate axes. Suppose the center is at $(h, k)$ and that the major axis is parallel to the $x$-axis

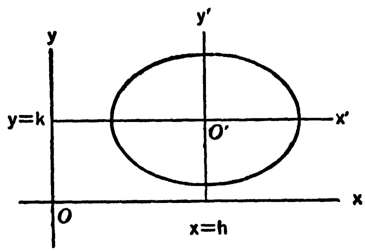



**Fig. 3-21**

(see Fig. 3-21). We use a new set of coordinate axes, parallel to the original set, with origin $O'$ at the center of the ellipse. If the new coordinates are $x'$, $y'$, the equation of the ellipse in the new system is

$$\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1.$$

But it is evident from the figure that

$$x' = x - h, \qquad y' = y - k. \tag{9}$$

Thus in the original coordinate system the equation of the ellipse is

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1. \tag{10}$$

If we carry out the squaring and collect the constant terms we get an equation of the form

$$Ax^2 + By^2 + Cx + Dy + E = 0, \tag{11}$$

in which $A$ and $B$ are both positive. If we have *any* equation of the form (11) (with $A$ and $B$ both positive) we can deal with it by completion of squares, much as we did in the case of the circle, and hence find out what the equation represents. As in the previous case there may be no graph at all, or the graph may consist of a single point. Otherwise the graph is a circle if $A = B \neq 0$, and an ellipse if $A$ and $B$ are both positive (or both negative) but unequal.

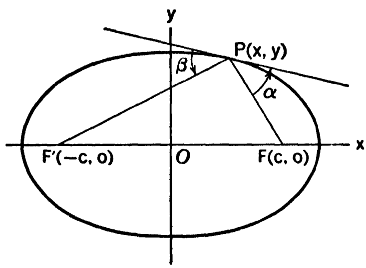



**Fig. 3-22**

One of the interesting things about an ellipse is the fact that the lines drawn from the two foci to a point on the ellipse make equal angles with the tangent at this point. This is illustrated in Fig. 3-22; the angles $\alpha$ and $\beta$ are equal. The proof is left for an exercise. This property of an ellipse lends itself to optical and acoustical applications.

## EXERCISES

**1.** Identify the graph (if any) of each equation. Draw the figure. If the graph is a circle, give the center and radius. If it is an ellipse, give the center, the foci, and the lengths of the major and minor axes.

(a) $x^2 + y^2 + 2x - 6y + 6 = 0.$

(b) $9x^2 + 4y^2 - 36x + 16y + 16 = 0.$

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

(d) $x^2 + y^2 + 4x + 2y + 6 = 0.$

(e) $9x^2 + 25y^2 - 50y = 200.$

(f) $x^2 + 2y^2 - 10x + 12y + 43 = 0$.
(g) $16x^2 + 25y^2 - 200x + 400 = 0$.
(h) $144x^2 + 144y^2 - 216x + 192y = 80$.
(i) $x^2 + y^2 + 14x - 10y + 10 = 0$.
(j) $9x^2 + 4y^2 + 18x - 16y + 12 = 0$.

**2.** Find the equation of the circle through the three given points.
(a) $(2, 2)$, $(2, -2)$, $(-4, 2)$.
(b) $(1, 6)$, $(2, 5)$, $(-6, -1)$.
(c) $(4, -2)$, $(2, 2)$, $(-5, 1)$.
(d) $(2, -3)$, $(5, -1)$, $(4, 3)$.

**3.** Find the equation of the circle:
(a) With center at $(-2, 3)$, and passing through $(1, -2)$.
(b) With center in the first quadrant on the line $x = 4$, and tangent to both axes.
(c) With center on the $x$-axis, and passing through $(2, 3)$ and $(6, 5)$.
(d) Having the points $(18, -4)$ and $(-6, 6)$ as the ends of a diameter.
(e) With center on the line $x - y + 1 = 0$, and passing through $(2, 1)$ and $(4, 3)$.
(f) With radius 6, center in the fourth quadrant, and passing through the points $(3, 2)$, $(-1, 0)$.
(g) Circumscribing the right triangle with vertices at $(1, 8)$, $(10, 5)$, and $(-2, -1)$.
(h) Through the mid-points of the sides of the triangle with vertices at $(-4, 0)$, $(2, 0)$, and $(0, 6)$.
(i) With center at $(6, 9)$, and tangent to the circle
$$x^2 + y^2 + 4x - 6y - 12 = 0.$$

**4.** Find the equation of the tangent to each circle or ellipse at the indicated point.
(a) $x^2 + y^2 = 169$ at $(5, 12)$.
(b) $(x - 3)^2 + (y + 2)^2 = 25$ at $(6, 2)$.
(c) $x^2 + 9y^2 = 225$ at $(9, 4)$.
(d) $x^2 + y^2 + 11x - 9y = 0$ at $(0, 0)$.
(e) $3x^2 + 3y^2 + 16x + 8y = 30$ at $(1, 1)$.
(f) $x^2 + 4y^2 - 2x + 8y = 35$ at $(3, 2)$.
(g) $5x^2 + 9y^2 - 10x - 54y = 63$ at $(2, -1)$.
(h) $9x^2 + 25y^2 - 50y = 200$ at $(5, 1)$.

**5.** Find the equation of the ellipse:
(a) With foci at $(\pm 4, 0)$ and major axis of length 12.
(b) With foci at $(0, \pm 5)$ and minor axis of length 16.
(c) With major and minor axes of lengths 5, 4, respectively, center at the origin, and foci on the $y$-axis.
(d) With foci at $(\pm 2, 0)$ and eccentricity $\frac{2}{5}$.
(e) With eccentricity $\frac{1}{2}$, center at the origin, and the ends of the major axis at $(0, \pm 8)$.
(f) With eccentricity $e = \frac{1}{10}$ and the ends of the minor axis at $(0, \pm 20)$.

6. Find the equation of the ellipse:
   (a) With ends of the major axis at $(-3, 2)$, $(5, 2)$ and 4 as the length of the minor axis.
   (b) With major axis 10 units long, and foci at $(5, 3)$ and $(1, 3)$.
   (c) With minor axis 8 units long and foci at $(1, -2)$ and $(1, 4)$.
   (d) With eccentricity $e = \frac{2}{3}$ and ends of the major axis at $(7, 1)$ and $(-5, 1)$.

7. Find the intersections of the circles $x^2 + y^2 + 2x - 14y + 25 = 0$, $x^2 + y^2 + x - 7y = 0$.

8. Find the length of the common chord of the circles $x^2 + y^2 + 6x - 8y = 1$, $x^2 + y^2 + 4x - 7y = 4$.

9. A point moves so that its distance from $(2, 2)$ is half its distance from $(-4, 3)$. Find the curve it describes.

10. A point moves so that the sum of the squares of its distances from $(3, 2)$ and $(-5, 2)$ is always 40. Find the curve it describes.

11. A point $P$ moves so that the distance from $(0, 0)$ to the mid-point of the line joining $P$ to $(3, 0)$ is always 4. Find the curve which $P$ describes.

12. A line segment of length 5 moves with one end $A$ on the $x$-axis and the other end $B$ on the $y$-axis. A point $P$ fixed on the segment is 3 units from $A$ and 2 units from $B$. Find the curve traced out by $P$ as the segment moves.

13. An ellipse has its center at the origin and its major axis along a coordinate axis. Find its equation if it goes through (a) $(4, 1)$ and $(2, 2)$, (b) $(-3, 1)$ and $(2, -4)$.

14. An ellipse has its center at the origin, its foci on the $x$-axis, and eccentricity $\frac{3}{5}$. Find its equation if it goes through $(12, 4)$.

15. Find the equation of the circle which is tangent to the ellipse $16x^2 + 25y^2 = 400$ at the points for which $x = 3$.

16. (a) Show that the slope of the tangent to the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ at $x = c$ $(y > 0)$ is $-e$. (b) Where does this tangent cut the $x$-axis?

17. Let $P$ be a point on the ellipse $b^2x^2 + a^2y^2 = a^2b^2$. Let $A$ be the point $(a, 0)$. Let $Q$ be the point of intersection of the line $x = -a$ and the tangent to the ellipse at $P$. Show that $OQ$ is parallel to $AP$.

18. (a) Refer to Fig. 3-22. Show that the slope of the tangent is $-b^2x/a^2y$. Using the formula for the tangent of the angle between two lines, show that $\tan \alpha = b^2/cy$. Show that this is also the value of $\tan \beta$. Hence $\alpha = \beta$.
    (b) Show that the equation of the tangent at $(x_0, y_0)$ is $\dfrac{x_0x}{a^2} + \dfrac{y_0y}{b^2} = 1$.

19. Two competing companies, $A$ situated at $(0, 40)$ and $B$ at $(30, 0)$ (units in miles), advertise to install equally priced furnaces in a buyer's house. Company $A$ adds a charge of 40 cents per mile (measured in a direct line) from its location to the house, while company $B$ adds a charge of 60 cents

per mile. In what region is it cheaper to have the furnace installed by company $B$?

## 3-9 Hyperbolas

A hyperbola may be defined as a curve consisting of all points $P$ in a plane such that the difference of the distances from $P$ to each of two given points $F'$ and $F$ (in the plane) is a constant. We denote the constant by $2a$. Then (see Fig. 3-23) either

$$F'P - FP = 2a \quad \text{or} \quad FP - F'P = 2a.$$

The curve is made up of two separate parts: the part on which $F'P > FP$, and the part on which $FP > F'P$. These parts are called *branches*. It is clear that the curve must be symmetric with respect to the line through $F'$ and $F$, and also with respect to the perpendicular bisector of the segment $F'F$. The points $F'$ and $F$ are called the *foci* of the hyperbola. The line

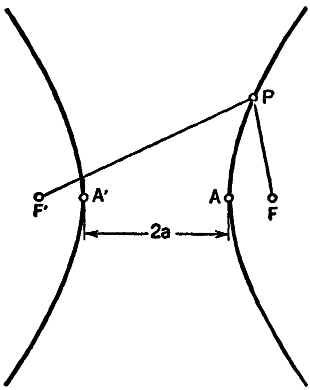



**Fig. 3-23**

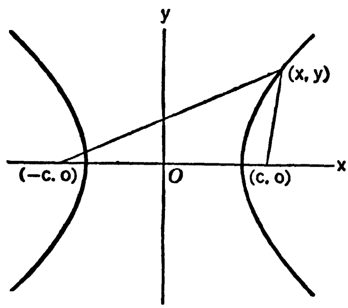



**Fig. 3-24**

through the foci is cut by the curve in points $A'$ and $A$, called the *vertices* of the hyperbola. It is easily seen from the definition that $A'A = 2a$, for

$$F'A - AF = 2a \quad \text{and} \quad F'A' = AF.$$

To obtain an equation for the hyperbola, place the foci on the $x$-axis, at the points $(\pm c, 0)$, equal distances on either side of the origin (see Fig. 3-24). The definition of the hyperbola is expressed by the two equations

$$\sqrt{(x + c)^2 + y^2} - \sqrt{(x - c)^2 + y^2} = \pm 2a \tag{1}$$

($+2a$ for the right branch, and $-2a$ for the left branch). We proceed to simplify this by squaring, just as we did in the case of the ellipse. The algebra is exactly the same, and we arrive at the equation

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2). \tag{2}$$

It can be shown that any point which satisfies (2) must satisfy one of the two equations (1), so (2) is an equation which describes the hyperbola.

It is convenient to define a positive number $b$ by the formula

$$b = \sqrt{c^2 - a^2}, \quad \text{or} \quad c^2 = a^2 + b^2. \tag{3}$$

We can do this, because $c > a$. The equation of the hyperbola then becomes $-b^2x^2 + a^2y^2 = -a^2b^2$, or

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{4}$$

The hyperbola has two asymptotes, which are the lines

$$y = \frac{b}{a}x \quad \text{and} \quad y = -\frac{b}{a}x. \tag{5}$$

The relation of the curve to these asymptotes is shown in Fig. 3-25. This figure also shows the relation of $a$, $b$, and $c$. The rectangle of dimensions $2a$ by $2b$ has the asymptotes as diagonals. The circle of radius $c$, circumscribed about the rectangle, passes through the foci.

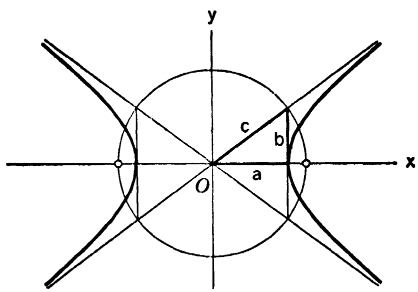



**Fig. 3-25**

In order to prove that the lines (5) actually are asymptotes of the hyperbola we proceed as follows. Let $(x_0, y_0)$ be a point on the hyperbola in the first quadrant. We shall show that, as $x_0 \to +\infty$, the tangent to the hyperbola at this point approaches coincidence with the line $y = bx/a$. In view of the symmetry of the hyperbola, this will be adequate proof of the assertion about the two asymptotes.

The slope of the hyperbola at any point is found by differentiation:

$$\frac{2x}{a^2} - \frac{2y}{b^2}\frac{dy}{dx} = 0, \qquad \frac{dy}{dx} = \frac{b^2x}{a^2y}.$$

Hence the equation of the tangent at $(x_0, y_0)$ is

$$y - y_0 = \frac{b^2x_0}{a^2y_0}(x - x_0). \tag{6}$$

We can simplify this by using the fact that $b^2x_0^2 - a^2y_0^2 = a^2b^2$ (which expresses the fact that $(x_0, y_0)$ is on the hyperbola). Clearing fractions in (6), we have

$$a^2y_0y - a^2y_0^2 = b^2x_0x - b^2x_0^2, \qquad b^2x_0x - a^2y_0y = b^2x_0^2 - a^2y_0^2 = a^2b^2,$$

or

$$\frac{x_0x}{a^2} - \frac{y_0y}{b^2} = 1. \tag{7}$$

This is the standard form of the equation of the tangent. The slope is $b^2x_0/a^2y_0$; the $y$-intercept is $-b^2/y_0$. Now, we are considering a point in the first quadrant, so when $x_0 \to +\infty$ we have $y_0 \to +\infty$ and $b^2/y_0 \to 0$. We must find the limiting value of the slope. Now

$$\frac{y_0^2}{b^2} = \frac{x_0^2}{a^2} - 1, \qquad y_0 = \frac{b}{a}\sqrt{x_0^2 - a^2},$$

$$\frac{b^2x_0}{a^2y_0} = \frac{b^2x_0}{ab\sqrt{x_0^2 - a^2}} = \frac{b}{a}\frac{x_0}{\sqrt{x_0^2 - a^2}}.$$

When $x_0 \to +\infty$,

$$\frac{x_0}{\sqrt{x_0^2 - a^2}} = \frac{1}{\sqrt{1 - \dfrac{a^2}{x_0^2}}} \to 1.$$

The limiting value of the slope is seen to be $b/a$. We have now shown that, when $x_0 \to +\infty$, the tangent $(x_0, y_0)$ approaches coincidence with the line $y = bx/a$, for the latter line has slope $b/a$ and $y$-intercept 0.

The ratio $c/a$ is called the eccentricity of the hyperbola, and denoted by $e$:

$$e = \frac{c}{a}. \tag{8}$$

Observe that $e > 1$. When $e$ is near 1, then $b$ is near 0 and the hyperbola lies in a very small angle between the asymptotes. When $e$ is large, this indicates that $c$ is much larger than $a$, and that $b$ also is much larger than $a$. The angle between the asymptotes is large, and the curve is rather flattened at the vertices.

**Example 1:** The foci of a hyperbola are at $(\pm 4, 0)$, and the asymptotes make 30° angles with the line through the foci. Find the equation of the curve.

We know that $c = 4$. We find $a$ and $b$ with the aid of the triangle in Fig. 3-26.

$$\frac{b}{4} = \sin 30° = \frac{1}{2}, \qquad b = 2,$$

$$a^2 = c^2 - b^2 = 16 - 4 = 12, \qquad a = 2\sqrt{3}.$$

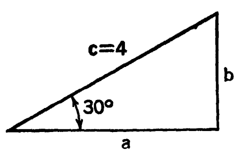


**Fig. 3-26**

The equation of the hyperbola is

$$\frac{x^2}{12} - \frac{y^2}{4} = 1.$$

If the hyperbola is placed with its foci on the $y$-axis and its center of symmetry at the origin, the equation is

$$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1, \quad \text{or} \quad \frac{x^2}{b^2} - \frac{y^2}{a^2} = -1. \tag{9}$$

The distinction between this equation and the equation when the foci are on the $x$-axis depends, not on the relative magnitudes of $a$ and $b$, but by the placing of the minus sign in the equation. With a hyperbola we can have $a > b$, $a = b$, or $a < b$. If $a = b$ the hyperbola is often called a *rectangular hyperbola,* because in this case the asymptotes intersect at right angles.

**Example 2:** Find the hyperbola with foci at $(\pm 2\sqrt{13}, 0)$ and the lines $3y = \pm 2x$ as asymptotes.

We see that $b/a = \frac{2}{3}$. Since $c^2 = a^2 + b^2 = 4(13) = 52$, we have

$$52 = a^2 + \frac{4}{9}a^2 = \frac{13}{9}a^2, \qquad a^2 = 36, \qquad b^2 = 52 - 36 = 16.$$

Hence $a = 6$, $b = 4$. The equation is

$$\frac{x^2}{36} - \frac{y^2}{16} = 1.$$

**Example 3:** Identify the hyperbola $9y^2 - 6x^2 = 36$ and give the values of the constants associated with it.

The equation can be put in the form

$$\frac{y^2}{4} - \frac{x^2}{6} = 1.$$

Hence the foci are on the $y$-axis, and $a^2 = 4$, $b^2 = 6$, $c^2 = 4 + 6 = 10$. Then

$$a = 2, \qquad b = \sqrt{6}, \qquad c = \sqrt{10}, \qquad e = \frac{\sqrt{10}}{2}.$$

Just as with the ellipse, we can easily write the equation of a hyperbola with center at $(h, k)$ and foci on a line parallel to a coordinate axis. For the case of foci on a line parallel to the $x$-axis the equation is

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1. \tag{10}$$

This leads us to consider equations of the form

$$Ax^2 + By^2 + Cx + Dy + E = 0,$$

where $A$ and $B$ are of opposite sign. This kind of equation will usually represent a hyperbola, but it may in special cases represent two straight lines. The procedure of identification is that of completing the square.

**Example 4:** Identify the graph of the equation

$$9x^2 - 16y^2 + 36x + 160y + E = 0$$

for various values of $E$.

We write

$$9(x^2 + 4x \qquad) - 16(y^2 - 10y \qquad) = -E,$$
$$9(x + 2)^2 - 16(y - 5)^2 = -E + 36 - 400 = -E - 364.$$

There are now three cases:

*Case 1.* $-E - 364 > 0$, e.g., $E = -400$, $-E - 364 = 36$. Then we have

$$\frac{(x + 2)^2}{4} - \frac{(y - 5)^2}{\frac{9}{4}} = 1.$$

Here we have a hyperbola with center at $(-2, 5)$, foci on a line parallel to the $x$-axis, and $a = 2$, $b = \frac{3}{2}$. Hence $c = \frac{5}{2}$, and the foci are at $(-2 \pm \frac{5}{2}, 5)$.

*Case 2.* $-E - 364 < 0$, e.g., $E = 36$, $-E - 364 = -400$. Then we have

$$\frac{(x + 2)^2}{\frac{400}{9}} - \frac{(y - 5)^2}{25} = -1.$$

Here we have a hyperbola with center at $(-2, 5)$, $a = 5$, $b = \frac{20}{3}$, $c = \frac{25}{3}$, and with foci at $(-2, 5 \pm \frac{25}{3})$ on a line parallel to the $y$-axis.

*Case 3.* $-E - 364 = 0$, $E = -364$. Then we have

$$9(x + 2)^2 - 16(y - 5)^2 = 0,$$
$$3(x + 2) = \pm 4(y - 5).$$

Here we have the two lines

$$3x + 4y - 14 = 0, \qquad 3x - 4y + 26 = 0,$$

which intersect at $(-2, 5)$. These lines form the graph in the third case. They are the same as the asymptotes to the hyperbolas in the first two cases.

In studying these cases the student should draw the graphs.

We frequently meet with the equations of rectangular hyperbolas having the coordinate axes as asymptotes. The equation in such cases is of the form

$$xy = \text{a constant } (\neq 0).$$

If the constant is positive, the branches of the hyperbola lie in the first and third quadrants. If the constant is negative, the branches are in the second and fourth quadrants. For examples see Fig. 3-27 and Fig. 3-28.

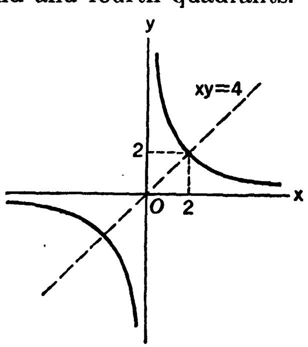



**Fig. 3-27**

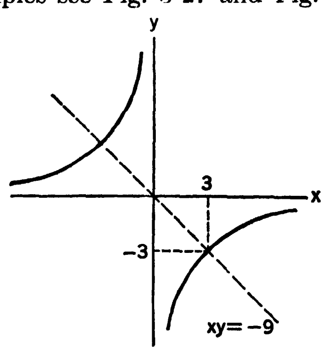



**Fig. 3-28**

To derive the equation of a rectangular hyperbola with the axes as asymptotes, let the foci be at $\left(\frac{c}{\sqrt{2}}, \frac{c}{\sqrt{2}}\right)$ and $\left(-\frac{c}{\sqrt{2}}, \frac{-c}{\sqrt{2}}\right)$. Since $a = b$ we have $c^2 = 2a^2$, $2a = c\sqrt{2}$. The definition of the hyperbola is expressed by the equations

$$\sqrt{\left(x + \frac{c}{\sqrt{2}}\right)^2 + \left(y + \frac{c}{\sqrt{2}}\right)^2} - \sqrt{\left(x - \frac{c}{\sqrt{2}}\right)^2 + \left(y - \frac{c}{\sqrt{2}}\right)^2} = \pm c\sqrt{2}.$$

On elimination of the radicals by squaring, just as we did in deriving (2) from (1), we arrive at the equation

$$xy = \frac{c^2}{4}. \tag{11}$$

The definition of a hyperbola makes this type of curve useful in various types of range-finding work. One example is that of locating an enemy artillery piece. Three range-finder listening posts are in contact by telephone. When the gun is fired, each post notes the time at which the shot is heard. By comparing with each other, each pair of posts can determine the difference in the distance from the gun to the two posts. This places the gun on a certain hyperbola with the two posts as foci. Using two different pairs of posts, two hyperbolas are found. The gun is then located graphically at the intersection of the hyperbolas.

Another use of hyperbolas is in blind flying. Two radar beacon stations are used as foci, and it is desired to make a plane fly a course following one branch of a hyperbola with foci at the beacons. Each station sends out a pulse signal which is picked up by the plane and registered on an instrument which shows the distance from the plane to the beacon. The plane then flies so as to maintain a prescribed constant difference in distance from the beacons. In practice one of the beacon signals is usually sent out with a preset delay, so that the plane maintains an *apparently* equal distance from the beacons.

There is a property of the hyperbola corresponding to the so-called "optical properties" of the parabola and ellipse. It is this: The tangent to a hyperbola at a point $P$ bisects the angle between the lines joining $P$ to the two foci. Proof of this is left for an exercise.

## EXERCISES

**1.** Draw the hyperbola in each case. Make a figure like that in Fig. 3-25, or a corresponding one if the foci are on the $y$-axis. Begin by finding $a$, $b$, $c$ and drawing the asymptotes.

(a) $\frac{x^2}{4} - \frac{y^2}{9} = 1.$

(b) $\frac{x^2}{25} - \frac{y^2}{4} = 1.$

(c) $\frac{x^2}{36} - \frac{y^2}{25} = 1.$

(d) $\frac{x^2}{16} - \frac{y^2}{100} = 1.$

**2.** Identify the graph and sketch it. Give all essential data as in the discussion of Example 4.
(a) $y^2 - 4x^2 - 4y - 40x - 116 = 0.$
(b) $4x^2 - 25y^2 + 24x + 50y + 11 = 0.$
(c) $x^2 - 25y^2 - 6x - 50y + 9 = 0.$
(d) $4x^2 - 9y^2 - 16x - 18y - 29 = 0.$
(e) $x^2 - y^2 + 14x + 14y = 49.$
(f) $21y^2 - 16x^2 + 42y + 96x = 459.$
(g) $16x^2 - 9y^2 - 64x - 72y = 656.$
(h) $4x^2 - 9y^2 - 24x + 18y + 27 = 0.$
(i) $4x^2 - 9y^2 - 24x + 18y + 63 = 0.$
(j) $9y^2 - 4x^2 - 18y + 24x + 27 = 0.$

**3.** Write the equation of the hyperbola with center at the origin, and:
(a) a focus at $(-10, 0)$, a vertex at $(6, 0)$;
(b) a vertex at $(5, 0)$, the line $5y = 4x$ an asymptote;
(c) focus at $(0, 4)$, $e = 2$;
(d) focus at $(-6, 0)$, $e = 3$;
(e) asymptotes $x \pm 2y = 0$, a vertex at $(3, 0)$;
(f) major axis from $(0, -\frac{3}{2})$ to $(0, \frac{3}{2})$, $e = 2$.

**4.** Find the equation of each hyperbola.
(a) Through $(6, 10)$, with $3y = \pm 4x$ as asymptotes.
(b) Through $(2, 3)$, with asymptotes $y = \pm 2x$.
(c) Through $(1, 3)$, with asymptotes $y = \pm 2x$.
(d) Through $(4, 3)$, with asymptotes $3y = \pm 2x$.
(e) Through $(-4, 2)$, with foci on the $x$-axis and $e = \sqrt{5}/4$.

**5.** Find the tangent to each hyperbola at the point indicated.
(a) $4x^2 - y^2 = 15$ at $(2, -1)$.
(b) $144x^2 - 25y^2 = 3800$ at $(13, \frac{144}{5})$.
(c) $9y^2 - 16x^2 = 324$ at $(-6, -10)$.
(d) $5x^2 - 4y^2 = 64$ at $(4, 2)$.
(e) $4x^2 - 25y^2 + 24x + 50y + 22 = 0$ at $(-1, 2)$.
(f) $5y^2 - 9x^2 + 10y + 54x = 112$ at $(7, 5)$.
(g) $3x^2 - y^2 + 12x + 8y = 7$, at $(0, 7)$.
(h) $xy = -64$ at $(4, -16)$.

**6.** Find the equation of the hyperbola with center at the origin and foci on one of the coordinate axes, if it goes through
(a) $(2, 3)$ and $(1, 2)$.
(b) $(4, 3)$ and $(5, 6)$.
(c) $(5, -3)$ and $(-1, 1)$.
(d) $(7, 2)$ and $(4, -1)$.

**7.** A point moves so that it is equidistant from the point $(c, 0)$ and the circle of radius $2a$ with center at $(-c, 0)$. What is the curve traced out by the point?

**8.** Show that all hyperbolas with the lines $y = \pm\frac{3}{4}x$ as asymptotes fall into two classes: (1) those with eccentricity $\frac{5}{4}$ and foci on the $x$-axis, and those with eccentricity $\frac{5}{3}$ and foci on the $y$-axis. Draw several hyperbolas of each type.

**9.** Find the intersections of the ellipse and the hyperbola:
(a) $x^2 + 4y^2 = 16$ and $16x^2 - y^2 = 16$.
(b) $x^2 + 3y^2 = 24$ and $3x^2 - y^2 = 12$.
In case (b) show that the foci of the ellipse are the same as the foci of the hyperbola, and that the curves intersect at right angles.

**10.** Consider the hyperbola $b^2x^2 - a^2y^2 = a^2b^2$, and the tangent to it at a point in the first quadrant. Show that the acute angle which this tangent makes with each of the lines from the foci to the point of tangency is the same, and that the tangent of this angle is $b^2/cy$, where $y$ is the ordinate of the point of tangency.

**11.** An ellipse and a hyperbola have the same foci. Prove that the curves intersect at right angles.

**12.** (a) Show that $\dfrac{x^2}{25 + k} + \dfrac{y^2}{16 + k} = 1$ represents an ellipse if $k > -16$ and a hyperbola if $-25 < k < -16$, and that all these curves have the same foci. (b) Find the eccentricity of each ellipse and each hyperbola in terms of $k$. (c) Discuss what happens to the eccentricity as $k \to +\infty$, as $k \to -16$, and as $k \to -25^+$. Draw the curves for $k = 11, 0, -15, -17, -20, -24$. (d) Find the intersections of the curves for $k = 0$ and $k = -20$, and the tangent to each curve at the first quadrant point of intersection.

**13.** In this problem the hyperbola $b^2x^2 - a^2y^2 = a^2b^2$ is considered.
(a) Find the tangent at the point of the curve in the first quadrant for which $x = c$. Show that its slope is $e$ and that it intersects the $x$-axis at $x = a/e$.
(b) Find the equation of the circle which is tangent to the curve at the two points for which $y > 0$ and $x = \pm c$.
(c) A tangent is drawn at a point $P$ on the curve in the first quadrant. A line $L$ is drawn through the focus $(c, 0)$ and perpendicular to the aforementioned tangent. Prove that $OP$ and $L$ intersect on the line $x = a/e$.
(d) If $D$ is the point in which the tangent in (c) intersects the line $x = a/e$, prove that $DFP$ is a right triangle with right angle at $F(c, 0)$.
(e) For the tangent described in (c), show that the point of tangency is midway between the points of intersection of the tangent with the asymptotes.
(f) Let $M$ and $N$ be the points where the tangent mentioned in (c) intersects the lines $x = -a$, $x = a$, respectively. Prove that the circle on $MN$ as diameter goes through the foci of the hyperbola.

**14.** Three listening posts are at $A(-2, 1)$, $B(3, 1)$, and $C(-2, 14)$ (1 unit = 0.1 mile). An enemy gun is fired, and the explosion is heard at $A$ 1.5 seconds after it is heard at $B$, and 2.5 seconds after it is heard at $C$. Take 0.2 mile per second as the speed of sound. Compute $a$, $b$, $c$ for the hyperbola with foci at $A$ and $B$ on which the gun must be located, and do the same for $A$ and $C$. Draw the asymptotes to these hyperbolas and locate the gun graphically, by assuming that the hyperbolas are indistinguishable from their asymptotes near the gun. Find the coordinates of the gun by finding algebraically the intersection of the appropriate pair of asymptotes.

## 3-10 Maxima and Minima

There are many interesting problems of the type: "when is such and such a thing the largest it can possibly be?" or "under what conditions does a certain variable quantity reach a minimum value?"

**Example 1:** If all isosceles triangles with perimeter 18 inches are considered, what are the dimensions of the triangle of greatest possible area?

**Example 2:** What positive number is such that the sum of the reciprocal of the number and four times its square is the smallest possible?

**Example 3:** A man is in a ploughed field, 300 feet from the nearest point $A$ of a straight road bordering the field. He wants to walk to a point $B$ on the road 600 feet from $A$. He can walk 3 feet per second in the ploughed field and 5 feet per second on the road. What is the least time in which he can walk to $B$?

The solution of such problems by calculus is an application of certain parts of the general theory of maxima and minima for functions of one independent variable. This theory is based on what we already know about the use of the first and second derivatives in curve tracing. We shall solve the problems posed in the foregoing examples, and use our solutions to illustrate statements about general theory and procedure.

**Example 1:** Denote the length of the base of the isosceles triangle by $2x$ (see Fig. 3-29). Since the perimeter is 18, the length of one of the equal sides is $\frac{1}{2}(18 - 2x) = 9 - x$. The altitude of the triangle is $\sqrt{(9 - x)^2 - x^2} = \sqrt{81 - 18x} = 3\sqrt{9 - 2x}$, and the area is

$$A = 3x\sqrt{9 - 2x}.$$

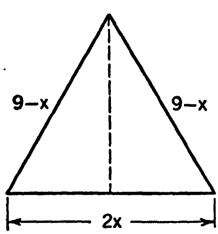



**Fig. 3-29**

The problem is: for what value of $x$ is $A$ largest? The admissible values of $x$ are from 0 to $\frac{9}{2}$. At the extremes $x = 0$ or $x = \frac{9}{2}$ we have $A = 0$; for such values of $x$ the triangle collapses into a line segment. We consider $A$ as a function of $x$ on the interval $0 \le x \le \frac{9}{2}$, and ask for the maximum value of $A$. Theorem 2-A guarantees that there is an $x$ for which $A$ is largest. Since $A = 0$ when $x = 0$ and when $x = \frac{9}{2}$, the maximum must occur at some point inside the interval. Theorem 2-B tells us that we must have $dA/dx = 0$ at the point where the maximum occurs. Hence we compute the derivative

$$\frac{dA}{dx} = 3x\frac{-2}{2\sqrt{9-2x}} + 3\sqrt{9-2x} = \frac{-3x+3(9-2x)}{\sqrt{9-2x}} = \frac{9(3-x)}{\sqrt{9-2x}}.$$

From this result we see that $dA/dx > 0$ if $0 \le x < 3$, $dA/dx = 0$ if $x = 3$, and $dA/dx < 0$ if $3 < x < \frac{9}{2}$. Hence $A$ increases when $x < 3$, reaches a maximum when $x = 3$, and decreases when $x > 3$. The triangle of maximum area is equilateral, the length of each side being 6.

**Example 2** (as stated at the beginning of this section): Denote the positive number by $x$. Then the sum of the reciprocal and four times the square is

$$y = \frac{1}{x} + 4x^2.$$

We wish to find $x$ so that $y$ is as small as possible. All positive values of $x$ are admissible. We compute the derivative:

$$\frac{dy}{dx} = -\frac{1}{x^2} + 8x = \frac{8x^3-1}{x^2}.$$

Now consider the behavior of $y$ as $x$ varies. When $x$ is near 0, $y$ is very large, and $dy/dx < 0$. As $x$ increases from near 0, $y$ decreases as long as $8x^3 - 1 < 0$. When $8x^3 - 1 = 0$, i.e., when $x = \frac{1}{2}$, $y$ reaches a minimum. When $x > \frac{1}{2}$, $dy/dx > 0$ and $y$ increases, becoming very large as $x$ gets large. Thus $x = \frac{1}{2}$ is the required value of $x$.

In the theory of maxima and minima we make a distinction between a *relative* maximum and an *absolute* maximum. A function is said to have a *two-sided relative maximum* at $x = x_0$ if there is *some* interval with $x_0$ at its center such that $f(x) \le f(x_0)$ for each $x$ in the interval. That is, $f(x_0)$ is the largest value of $f(x)$ when $x$ is restricted to lie in some interval (perhaps quite a short interval) extending on either side of $x_0$. A two-sided relative minimum is defined in a similar manner; we then require $f(x_0) \le f(x)$ instead of $f(x) \le f(x_0)$. When we speak of an *absolute* maximum, we always have in mind a definite collection of admissible values of the independent variable $x$, and this extent of variation of $x$ is specified ahead of time. The absolute maximum is the largest value of $f(x)$ when all the admissible values of $x$ are taken into account. Likewise for an absolute minimum.

In a given problem a two-sided relative maximum *may* also be an absolute maximum, but it need not be so in every problem. If an absolute maximum occurs when $x$ is inside (not at an end) of an interval of admissible values of $x$, then we also have a two-sided relative maximum. But if the absolute maximum occurs at one end of the interval, it is not a two-sided relative maximum. Suppose the admissible values of $x$ are those for which $a \le x \le b$. It is conceivable that there may be several two-sided relative maxima and minima for $f(x)$. Figure 3-30 illustrates a case in which $f(x)$ has two-sided relative maxima at $x_2$ and $x_4$, two-sided relative minima at $x_1$ and $x_3$, the absolute maximum at $x = a$, and the absolute minimum at $x_3$.

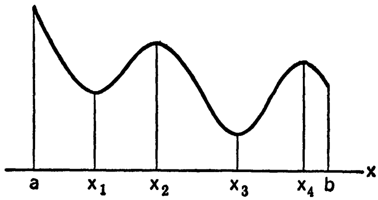



**Fig. 3-30**

When the function is differentiable for all admissible values of $x$, we know that $f'(x) = 0$ at each two-sided relative extreme. This is so by Theorem 2-B. Of course, there may be a point where $f'(x) = 0$ which is neither a relative maximum nor a relative minimum (e.g., $x = 0$ if $f(x) = x^3$). But we do have the following test.

THEOREM 3-G. *Suppose $f'(x_0) = 0$, that $f'(x) > 0$ if $x$ is near $x_0$ on the left, and that $f'(x) < 0$ if $x$ is near $x_0$ on the right. Then $f(x)$ attains a two-sided relative maximum at $x_0$. Likewise, if $f'(x_0) = 0$, $f'(x) < 0$ when $x$ is near $x_0$ on the left, and $f'(x) > 0$ when $x$ is near $x_0$ on the right, $f(x)$ attains a two-sided relative minimum at $x_0$.*

*Proof.* In the first case $f(x)$ increases with $x$ when $x$ is near $x_0$ on the left, and decreases as $x$ increases beyond $x_0$ on the right. Thus there is a relative maximum at $x_0$. The law of the mean exhibits the situation very clearly. If $x$ is near $x_0$ but distinct from it, the law of the mean asserts that

$$f(x) - f(x_0) = (x - x_0)f'(X),$$

where $X$ is some number between $x_0$ and $x$. If $x < x_0$, $f'(X) > 0$, and so $f(x) - f(x_0) < 0$, or $f(x) < f(x_0)$. If $x_0 < x$, $f'(X) < 0$, and once more we have $f(x) < f(x_0)$, because $x - x_0$ and $f'(X)$ are of opposite sign. Thus the value $f(x_0)$ is a relative maximum. In the second case $x - x_0$ and $f'(X)$ are of the same sign, so that $f(x) > f(x_0)$, and we have a relative minimum at $x_0$.

If we do not wish to examine the sign of the first derivative on either side of $x_0$, we can use the second derivative to distinguish between a relative maximum and a relative minimum.

THEOREM 3-H. *Suppose $f$ is differentiable on an interval, that $x_0$ is a point of the interval, not at either end, and that $f$ has a second derivative at $x_0$. Then $y = f(x)$ will have a two-sided relative maximum at $x_0$ if*

$$\frac{dy}{dx} = 0 \quad \textit{and} \quad \frac{d^2y}{dx^2} < 0 \quad \textit{at} \quad x_0.$$

*The function will have a two-sided relative minimum at $x_0$ if*

$$\frac{dy}{dx} = 0 \quad and \quad \frac{d^2y}{dx^2} > 0 \quad at \quad x_0.$$

*Proof.* In the first case the fact that $f'(x_0) = 0$ and $f''(x_0) < 0$ means that

$$\lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f'(x)}{x - x_0} < 0.$$

Hence $f'(x)$ is opposite in sign to $x - x_0$ when $x$ is near $x_0$. This means $f'(x) > 0$ if $x < x_0$ and $f'(x) < 0$ if $x > x_0$. As we saw in Theorem 3-G, these conditions indicate a relative maximum at $x_0$. The argument in the second case, for a relative minimum, is similar.

When we are looking for an absolute maximum (or minimum), we may proceed as follows: (1) Formulate the problem in a functional form, using some convenient independent variable. (2) Determine the admissible values of the independent variable. (3) Compute the derivative of the function and locate the two-sided relative maxima and minima of the function. (4) Decide whether the absolute extreme occurs at one of the points found in (3), or whether it occurs at an end point of an interval of admissible values.

In many problems there is only one two-sided relative extreme of the function, and this coincides with the absolute extreme. This is the case in Examples 1 and 2. In each of these cases it is clear that the absolute extreme does not occur at an end of an interval of admissible values.

We now turn to a study of the third example.

**Example 3** (as stated at the beginning of this section): In order to reach $B$ in the least possible time the man should walk across the field to some point $P$ on the road between $A$ and $B$ (see Fig. 3-31), and then walk to $B$. At the outset

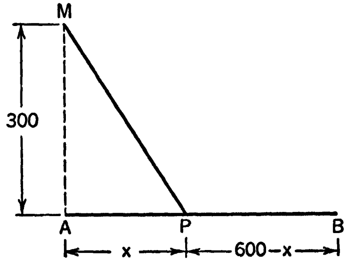



**Fig. 3-31**

it is conceivable that it may be best for $P$ to coincide with $A$ or $B$. We denote the distance $AP$ by $x$. The time to walk a certain distance is the distance divided by the rate of walking. Hence the total time $T$ (in seconds) required for the man to reach $B$ is

$$T = \frac{\sqrt{(300)^2 + x^2}}{3} + \frac{600 - x}{5}.$$

We wish to choose $x$ so that $T$ is an absolute minimum. The interval of admissible values of $x$ is $0 \le x \le 600$. At the ends of the interval we have

$$x = 0, \qquad T = 220,$$
$$x = 600, \qquad T = 100\sqrt{5} = 223.61.$$

The derivative is

$$\frac{dT}{dx} = \frac{x}{3\sqrt{(300)^2 + x^2}} - \frac{1}{5} = \frac{5x - 3\sqrt{(300)^2 + x^2}}{15\sqrt{(300)^2 + x^2}}.$$

This derivative is negative when $x = 0$, which indicates that, for $x$ positive and near 0, $T$ is less than 220. Solving to find when $dT/dx = 0$, we have

$$5x = 3\sqrt{(300)^2 + x^2}, \qquad 25x^2 = 9[(300)^2 + x^2],$$
$$16x^2 = 9(300)^2, \qquad x = \frac{900}{4} = 225.$$

The other solution, $x = -225$, is rejected because it is not in the interval of admissible values. The value of $T$ corresponding to $x = 225$ is $T = 200$. This is the absolute minimum. We know that this is so because $T$ has only one two-sided relative minimum (namely $T = 200$), and this value is smaller than the values of $T$ at the ends $x = 0$, $x = 600$.

## EXERCISES

1. A rectangular box has a square base and no top. The combined area of the sides and bottom is 48 square feet. Find the dimensions of the box of maximum volume meeting these specifications.

2. If in Exercise 1 the box is required to contain 108 cubic feet, find the dimensions which will give it the least total area of sides and bottom.

3. A farmer wishes to fence off a rectangular pasture along a straight river, the side along the river requiring no fence. He has barbed wire enough to build a fence one mile long. What is the area of the largest pasture of the above description which he can fence?

4. A triangle of base $b$ and altitude $a$ has acute base angles. A rectangle is fitted inside the triangle, one side resting on the base of the triangle. Show that the maximum possible area of the rectangle is half the area of the triangle.

5. Express the number 4 as the sum of two positive numbers in such a way that the sum of the square of the first and the cube of the second is as small as possible.

6. A rectangle is required to have a fixed perimeter $P$. Show that the rectangle of greatest area is a square.

7. Express the number 12 as the sum of two positive numbers in such a way that the product of one by the square of the other is as large as possible.

**8.** Find the positive number such that the sum of its square and 16 times the square of its reciprocal is as small as possible.

**9.** A tent-shaped solid has a square base and equal isosceles triangles at the ends, as in Fig. 3-32. If the perimeter of one of the triangular ends is 25 feet, find the dimensions of the figure so as to give it the maximum volume.

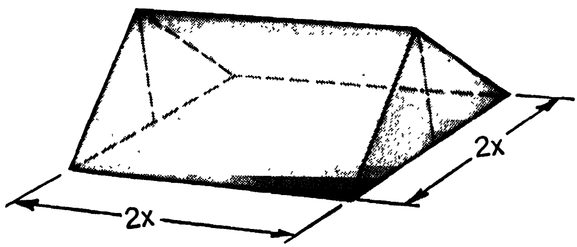



**Fig. 3-32**

**10.** (a) Suppose, in Example 3 of the text, that the point $B$ had been 200 feet from $A$. In that case what should the man have done in order to walk to $B$ in the least possible time?
(b) Modify Example 3 of the text by supposing the man initially $m$ feet from the road, but leave the other details unchanged. Find $x$ in terms of $m$ so as to make $T$ a minimum. What condition on $m$ makes it quickest for the man to walk directly to $B$ (i.e., $x = 600$)?

**11.** A house at $A$ is in the woods 12 miles north of an east-west road, the nearest point of which is $B$. At $C$, 5 miles east of $B$ on the road, is an electric power substation. If a power line is built from $C$ to $A$, it costs $r$ times as much per mile to build it through the woods as along the highway ($r$ is fixed). The line will either be built directly from $C$ to $A$, or along the road to a point $P$ part way toward $B$, and then through the woods to $A$, whichever is cheaper. Examine the situation for cheapest cost of construction (a) if $r = 3$; (b) if $r = 2$. (c) Find the largest value of $r$ for which it is cheapest to build the line directly from $C$ to $A$.

**12.** A circular ring of radius $b$ is uniformly charged with electricity, the total charge being $Q$. The force exerted by this charge on a unit particle $x$ units from the center of the ring, in a direction perpendicular to the plane of the ring, is $F = Qx(x^2 + b^2)^{-3/2}$. Examine the way in which $F$ varies as $x$ varies over all positive values. Find the absolute maximum of $F$ for such $x$, and explain clearly how you know you have not found a minimum or a relative, but not absolute, maximum.

**13.** A rectangle is to have an area of 64 square inches. Find its dimensions so that the distance from one corner to the mid-point of a nonadjacent side shall be a minimum.

**14.** At noon ship A, steaming east at 16 miles an hour, is due south of ship B which is steaming south at 12 miles an hour. They are 100 miles apart at noon. At what time are they closest together, and what is the distance between them then?

**15.** The maker of a certain article finds that, in order to sell $x$ of the articles each week, he must price them at $\$\sqrt{900 - 2x^2}$ apiece. What number of articles per week will bring him the greatest total revenue?

**16.** In a certain manufacturing process a plant produces $\dfrac{25 - 2x}{15 - x}$ tons per day of a high-grade product as an adjunct of the production of $x$ tons per day of a low-grade product. In order to operate at all, the plant must produce at least one ton per day of the low-grade product and the capacity production is $12\frac{1}{2}$ tons per day of this product. If the high-grade product brings $\frac{9}{5}$ as much per ton as the low-grade product, find the daily output of the low-grade product which will maximize the total revenue.

**17.** The rate $Q$ at which water flows over a certain spillway is proportional to $D(H - D)^{1/2}$, where $D$ is the depth of the flow and $H$ is the head. For a fixed value of $H$, what value of $D$ makes $Q$ a maximum?

**18.** A right circular cylinder, radius of base $r$, is inscribed in a right circular cone, radius of base $R$ and altitude $H$. Show that the volume of the cylinder is largest if $r = \frac{2}{3}R$. For what value of $r$ is the lateral area of the cylinder greatest?

**19.** Two points $A$ and $B$ are situated on the same side of a straight line $L$. Show that if $P$ is on $L$, the sum of the distances $AP$ and $PB$ is shortest

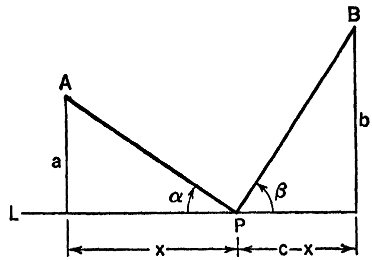



**Fig. 3-33**

when the angles $\alpha$ and $\beta$ are equal (see Fig. 3-33). Use the distance $x$ as independent variable and interpret the condition for minimum distance in terms of the cosines of $\alpha$ and $\beta$.

**20.** Two points $A$ and $B$ are on opposite sides of a straight line $L$. A particle is required to travel from $A$ to a point $P$ on $L$ at the speed $v_1$, and from $P$ to $B$ at a speed $v_2$. When $P$ is located so that the total traveling time from $A$ to $B$ is least, the acute angles $\theta_1$ and $\theta_2$ which the lines $AP$ and $PB$, respectively, make with the line perpendicular to $L$ at $P$ are such that

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1}{v_2}.$$

Prove this, using a procedure somewhat like that suggested in Exercise 19. Since a ray of light always travels from one point to another in the least possible time, the result of this exercise expresses the fundamental law of refraction (Snell's law) in optics, for the case of a ray of light passing from one homogeneous medium into another. For example, if from $A$ to $P$ is in air and from $P$ to $B$ is in water, $v_1 > v_2$, and so $\theta_1 > \theta_2$, the exact relation being that expressed in the foregoing formula.

**21.** A right circular cone of altitude $x$ is inscribed in a sphere of radius $R$. Show that when the cone has the greatest possible volume its altitude is $4R/3$ and its volume is $\frac{8}{27}$ that of the sphere. Begin by expressing the radius of the base of the cone as a function of $x$.

**22.** The cost of fuel for running a certain river steamer at a speed of $v$ miles per hour in still water is $\$v^3/32$ per hour. Other operational costs are \$160 per hour. It is desired to make a trip to a certain town upstream, against a current of 4 miles per hour. Find the most economical speed at which to make the trip.

**23.** An isosceles triangle is circumscribed about a circle of radius $R$. (a) Express one half the base of the triangle as a function of its altitude $x$. (b) Show that the area $A$ of the triangle is least when $x = 3R$. Suggestion: $A$ is least when $A^2$ is least. (c) Revolve the figure about the altitude of the triangle and so generate a right circular cone circumscribed about a sphere of radius $R$. Show that the least possible volume of the cone is twice the volume of the sphere. (d) Show that the lateral area of the cone is least when $x = R(2 + \sqrt{2})$.

**24.** A long sheet of paper is $c$ units wide. One corner of the paper is folded over as shown in Fig. 3-34. Find the value of $x$ which gives the triangle $ABC$ the least possible area.

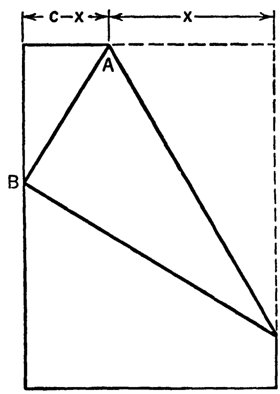



**Fig. 3-34**

**25.** A cylindrical hole of radius $x$ is bored through a sphere of radius $R$, the axis of the hole passing through the center of the sphere. Find $x$ so that the complete surface area of the remaining solid is as large as possible, and show that this area is $3\sqrt{3}/4$ times the area of the complete sphere.

## 3-11 Extremal Problems with Side Conditions

In many maximum or minimum problems the quantity to be made an extreme is naturally expressed as a function of two variables, and these two variables are related by an equation which expresses some condition inherent in the problem. Such a condition is often called a *side condition.*

**Example 1:** A rectangle is inscribed in the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{1}$$

in the manner shown in Fig. 3-35. What are the dimensions of the rectangle when its area is the greatest possible?

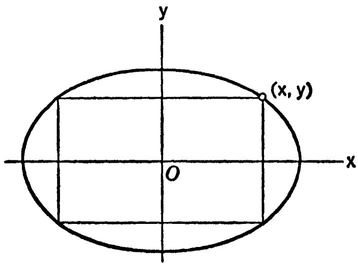



**Fig. 3-35**

If $(x, y)$ is the corner of the rectangle in the first quadrant, the area of the rectangle is $A = 4xy$. The relation between $x$ and $y$ is that expressed by (1). Now, one procedure for solving our problem would be to express $y$ in terms of $x$ and hence express $A$ as a function of $x$:

$$y = \frac{b}{a}\sqrt{a^2 - x^2}, \qquad A = \frac{4b}{a}x\sqrt{a^2 - x^2}.$$

We could then proceed by the methods of § 3-10 to find the value of $x$ which makes $A$ a maximum. Since $A^2$ and $A$ are maximized for the same value of $x$, we can even avoid the radical sign by writing $A^2 = \frac{16b^2}{a^2}x^2(a^2 - x^2)$. We wish, however, to illustrate a different procedure. We select one of the two variables, say $x$, as independent, and we regard the other variable, $y$, as being a function of $x$ determined implicitly by the side condition (1). We then calculate $dA/dx$ and set it equal to 0, since we wish $A$ to be a maximum:

$$\frac{dA}{dx} = 4x\frac{dy}{dx} + 4y = 0. \tag{2}$$

We also differentiate each term of the equation expressing the side condition:

$$\frac{2x}{a^2} + \frac{2y}{b^2}\frac{dy}{dx} = 0. \tag{3}$$

Next we eliminate $dy/dx$ between (2) and (3):

$$\frac{dy}{dx} = -\frac{y}{x} \quad \text{from (2),}$$

$$\frac{2x}{a^2} + \frac{2y}{b^2}\left(-\frac{y}{x}\right) = 0 \quad \text{[substituting in (3)].}$$

Simplifying, we have

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0, \quad \text{or} \quad y = \frac{b}{a}x.$$

Finally, we combine this result with (1) to find the values of $x$ and $y$:

$$\frac{x^2}{a^2} + \frac{x^2}{a^2} = 1, \quad x^2 = \frac{a^2}{2}, \quad x = \frac{a\sqrt{2}}{2}.$$

$$y = \frac{b\sqrt{2}}{2}.$$

The rectangle of maximum area has dimensions $a\sqrt{2}$ by $b\sqrt{2}$.

The method just illustrated (which we may call the *implicit function method*) is sometimes advantageous for the avoidance of algebraic complications which may arise if we attempt to eliminate one of the variables before doing the differentiation. Also, if the problem is one in which there is some kind of symmetry with respect to the variables which occur, this new method tends to maintain this symmetry in a useful way.

The procedure in the implicit function method is the same for a minimum problem as for a maximum problem. How then can we know, in a given problem, whether we are minimizing or maximizing the quantity we are examining? If there is real doubt, we may be forced to go back to the method of § 3-10, expressing everything in terms of one independent variable, at least for the purpose of settling the issue as to whether we are dealing with a maximum or a minimum. But in practice it frequently occurs that we can tell what to expect from the physical or geometrical nature of the problem. For instance, in Example 1, it is quite evident that the extremal problem for the area of the rectangle inscribed in the ellipse is a maximum problem, not a minimum problem. As to the actual existence of the maximum for some value of $x$ in the interval $0 \leq x \leq a$, we can appeal to Theorem 2-A, since it is clear that $y$, and hence the area, is a continuous function of $x$.

If we accept it as known that a given problem is genuinely a maximum problem, and if the implicit function method gives us just one answer, this answer is the solution to our problem. Likewise for a minimum problem. Of course, the method depends on certain assumptions about the implicit function defined by the side condition. However, this is not the place for a theoretical discussion of these assumptions nor of what might go wrong in exceptional cases.

**Example 2:** A cylindrical can without a top is to be formed from aluminum sheeting of uniform thickness, and is to weigh just $\frac{1}{4}$ pound. Find the relation between its height and the radius of the base when the volume of the can is greatest.

Let the height and radius be $h$ and $r$. The volume of the can is $V = \pi r^2 h$, and the area of the aluminum forming it is

$$A = \pi r^2 + 2\pi rh.$$

Here the side condition is expressed by the requirement that $A$ be a certain constant, determined by the weight of a square unit of the sheeting and the fact that the can must weigh a specified amount. So $A$ is constant and $V$ is variable. We select either $r$ or $h$ as independent. Let us choose $r$. Then, for maximum $V$ we want

$$0 = \frac{dV}{dr} = \pi\left(r^2\frac{dh}{dr} + 2rh\right). \tag{4}$$

Since $A$ is constant,

$$0 = \frac{dA}{dr} = \pi\left(2r + 2r\frac{dh}{dr} + 2h\right). \tag{5}$$

From (4) we have $dh/dr = -2h/r$. Combining this with (5) we obtain

$$r + r\left(\frac{-2h}{r}\right) + h = 0, \quad \text{or} \quad r = h.$$

This solves the problem as it was posed. The can should have a depth equal to the radius of the base.

## EXERCISES

**1.** Find the ratio of height to radius of base for a quart tin can with a top if the total surface area is the least possible.

**2.** A right circular cylinder is inscribed in a sphere of fixed radius. (a) Show that the cylinder has maximum volume when the diameter of its base is $\sqrt{2}$ times its altitude. (b) Show that the maximum lateral area which the cylinder can have is half the surface area of the sphere.

**3.** Find the dimensions of the rectangle of Example 1 when its perimeter is as large as possible. Show that the maximum perimeter is $4\sqrt{a^2 + b^2}$.

**4.** What ratio of height to radius of base will yield a right circular cone of greatest volume for a specified total surface area?

**5.** A certain dormer window is a rectangle surmounted by an equilateral triangle (with base the width of the rectangle). For a given area of the window opening, find the ratio of the height of the rectangle to its width, so as to minimize the perimeter of the window.

**6.** If in Exercise 5 the upper part of the window is a semicircle instead of a triangle, find the ratio of height to width so as to maximize the area of the window when the perimeter is specified.

**7.** A solid is formed by cutting hemispherical cavities from the ends of a right circular cylinder, the bases of the hemispheres coinciding with the ends of the cylinder. If the total area of the solid is a specified constant, find the ratio of height to radius of base for the cylinder so as to give the solid a maximum volume.

**8.** A north-south and an east-west road intersect at $C$. A diagonal road is to be constructed from a point $A$ north of $C$ to a point $B$ east of $C$, passing through a point $a$ miles east and $b$ miles north from $C$. Find the ratio $AC$ to $CB$ if the triangular area $ACB$ is the least possible.

**9.** An isosceles triangle is inscribed in the ellipse $b^2x^2 + a^2y^2 = a^2b^2$, with its vertex at $(a, 0)$. Find its altitude if the area is the maximum possible.

**10.** Find the shortest distance from the point (3, 0) to the hyperbola $y^2 - x^2 = 18$.

**11.** From a point $P_1(x_1, y_1)$ not on the ellipse $4x^2 + 9y^2 = 36$ a straight line is drawn to a variable point $P$ on the ellipse. Use the implicit function method to prove that when the distance $P_1P$ is a maximum or minimum, the line $P_1P$ is normal to the ellipse at $P$. Does your argument depend on the fact that you are dealing with this particular curve, or does it work for any curve?

**12.** A pyramid is to be constructed of plywood, with a square base and four equal triangular sloping faces. For a given total area of plywood, show that the volume of the pyramid is greatest when the height is $\sqrt{2}$ times the width of the base.

**13.** A tangent is drawn to the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ at a point $(x, y)$ in the first quadrant. (a) Find $y/x$ if the segment of the tangent cut off between the axes is as short as possible. (b) Find the length of the shortest segment.

**14.** The curve $5x^2 + 4xy + 2y^2 = 36$ is an ellipse with center at the origin, but it is twisted so that the major axis is not along a coordinate axis. Find the maximum and minimum distance from the origin to a point on the ellipse, and thus find the lengths of the major and minor axes of the ellipse. In doing this you will locate the ends of the major and minor axes.

## 3-12 Related Rates

It frequently happens, in problems of physical or geometrical interest, that two quantities are so related that each can be regarded as a function of the other, and that both are functions of the time. In such a situation, if we know the rate of change of one of the quantities with respect to time, we can find the rate of change of the other quantity without the need to know explicitly how either quantity depends on time.

**Example 1:** Suppose a cone of height 12 inches is changing in shape through the change of the radius of the base. What rate of increase of the radius will make the lateral area of the cone increase at the rate of $10\pi$ square inches per minute when the radius of the base is 5 inches?

Denote the lateral area by $S$, the radius of the base by $r$. Then (in square inches)

$$S = \pi r\sqrt{r^2 + 144}.$$

Consequently, using the composite function rule, we have

$$\frac{dS}{dt} = \pi r \frac{2r}{2\sqrt{r^2 + 144}} \frac{dr}{dt} + \pi\sqrt{r^2 + 144}\, \frac{dr}{dt}.$$

We put $dS/dt = 10\pi$, $r = 5$, and solve for $dr/dt$:

$$10\pi = 5\pi \frac{5}{13} \frac{dr}{dt} + 13\pi \frac{dr}{dt},$$

$$\frac{dr}{dt} = \frac{130}{194} = \frac{65}{97} = 0.67 \quad \text{(inch per second).}$$

In working rate problems we must recall that a negative rate of change implies a decreasing quantity.

**Example 2:** In a right triangle with hypotenuse of constant length 15 feet, one side is increasing in length while the other side decreases. If at a certain instant one side is 9 feet long and is increasing 4 inches per second, find the rate of change of the other side at this same instant.

Here we denote the lengths of the sides by $x$ and $y$, so that

$$x^2 + y^2 = 15^2 = 225.$$

Then
$$2x \frac{dx}{dt} + 2y \frac{dy}{dt} = 0, \qquad \frac{dy}{dt} = -\frac{x}{y} \frac{dx}{dt}.$$

We may suppose that $x$ refers to the increasing side. We put in $x = 9$, and compute

$$y = \sqrt{225 - 81} = 12.$$

Since 4 inches $= \frac{1}{3}$ foot, we have $dx/dt = \frac{1}{3}$. Then

$$\frac{dy}{dt} = -\frac{9}{12}\left(\frac{1}{3}\right) = -\frac{1}{4}.$$

The longer side is decreasing $\frac{1}{4}$ foot (or 3 inches) per second when it is 12 feet long.

Similar methods apply when more than two quantities are related. For example, the area of a rectangle depends on its length and breadth. If the rate of change of any two of the three quantities is known, the rate of change of the third quantity can be computed in terms of the dimensions of the rectangle.

## EXERCISES

1. A guy wire is to pass from the top of a pole 40 feet high to an anchorage on the ground 30 feet from the base of the pole. One end of the wire is made fast to the anchorage, and a man climbs the pole with the wire, keeping it taut. If he climbs $1\frac{1}{2}$ feet per second, how fast is he paying out the wire when he reaches the top of the pole?

2. A ladder 15 feet long rests against a house. It slides down, the lower end slipping along the level ground at the rate of 2 feet per second. How fast is the upper end of the ladder sliding down the wall when it is 12 feet from the ground?

3. A bomber is in level flight at 8 miles above the ground. The flight path passes directly over a rocket installation. How fast is the bomber flying if the airline distance to the rocket installation is decreasing at 4 miles per minute and this distance is 10 miles?

4. The area of a rectangle is increasing at the rate of 16 square inches per second. If one side is 12 inches and is increasing 5 inches per second, how fast is the other side changing when it is 8 inches?

5. The volume of a cylinder is increasing at the rate of $4\pi$ cubic centimeters per second. The radius of the base is increasing at the rate of 2 centimeters per second. How fast is the height of the cylinder changing when the volume is $36\pi$ cubic centimeters and the radius of the base is 3 centimeters?

6. Two airplanes fly eastward on parallel courses 12 miles apart. One flies at 240 miles per hour, the other at 300 miles per hour. How fast is the distance between the planes changing when the slower plane is 5 miles farther east than the faster plane?

7. Water is leaking through a hole in the vertex of a conical reservoir (vertex downward) at the rate of $24\pi$ cubic feet per minute. If the reservoir is 20 feet deep and 30 feet across the top, how fast is the depth of the water changing when the reservoir is $\frac{1}{8}$ full?

8. One ship is steaming at 10 knots straight north toward a port. Another ship is steaming at 15 knots on a course 30° south of east, directly away from the port. Find the rate of change of the distance between the ships when their distances from the port are, respectively, (a) 120 and 105 nautical miles; (b) 130 and 90 nautical miles; (c) 100 and 145 nautical miles. What is the special significance of the answer in (a)? *Note:* a knot is a speed of one nautical mile per hour.

9. A man is running over a bridge at a rate of 10 feet per second while a boat passes under the bridge and immediately below him at the rate of 20 feet per second. The boat's course is at right angles to the course of the man, and 20 feet below it. How fast are boat and man separating 1 second later?

10. A ladder 10 feet long is leaning against a wall 8 feet high, with its upper end projecting over the wall. If the lower end of the ladder slides away from the wall (on horizontal ground) at the rate of 2 feet per second, find the rate at which the upper end of the ladder is approaching the ground: (a) when 1 foot of the ladder is projecting over the wall; (b) when the top of the ladder reaches the wall.

*CHAPTER IV*

# TRIGONOMETRIC AND INVERSE TRIGONOMETRIC FUNCTIONS

## 4-1 Trigonometric Functions

The study of trigonometry is carried on with different aims at different levels of mathematical study. Our present interest is not primarily geometrical, but analytical. We wish to study the sine, cosine, tangent, and their reciprocals as *functions*. The sine function is a function which correlates with each angle a number called its sine. By selecting a particular system for measuring angles (i.e., by choosing a unit of measurement), the sine function becomes a function in the sense of § 1.6; that is, the numerical measure of the angle is the independent variable and the sine of the angle is the dependent variable. Actually, there is a different sine function for each choice of the unit of angular measurement. We do not know what sin 10 means until we know whether the 10 means 10 degrees, 10 radians, or 10 units of some other kind.

There are only two systems of angular measurement in common use: the *degree system* and the *radian system.* It is customary in elementary trigonometry and in a good deal of analytic geometry to use the degree system. But in calculus it is the standard practice to use one radian as the unit of angular measurement. This is purely for convenience, and we shall see why it is convenient when we learn how to differentiate the sine function.

The radian measure $\theta$ of an angle is defined by placing the vertex of the angle at the center of a circle (see Fig. 4-1), and taking $\theta$ to be the ratio of the intercepted arc to the radius:

$$\theta = \frac{s}{r}.$$

One radian is that angle for which $\theta = 1$, and hence $s = r$. If $r = 1$, note that $\theta = s$. The use of radian measure depends upon knowledge about the lengths of arcs of a circle. In particular, we need to know that the circumference of a circle of radius $r$ is $2\pi r$.

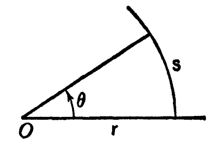


Fig. 4-1

If $\phi$ is the measure of an angle in degrees and $\theta$ is the measure in radians, $\phi$ and $\theta$ are proportional, so that $\theta = k\phi$, where $k$ is a constant. The value of $k$ is found by inserting a pair of corresponding values of $\theta$ and $\phi$. If $\phi = 180°$, the arc subtended by the angle is a semicircle, so that $s = \pi r$, and $\theta = \pi$. Thus $\pi = 180k$. This gives $k = \pi/180$, so that we have the general formula

$$\theta = \frac{\pi}{180}\phi$$

connecting $\theta$ and $\phi$. In particular, if $\theta = 1$, $\phi = 180/\pi = 57.2957\ldots$, so that an angle of 1 radian contains approximately 57.3 degrees.

For the definitions of the trigonometric functions we refer back to the trigonometry review at the end of § 1.3. From now on, however, in all references to trigonometric functions we shall assume that $\sin\theta$ denotes the sine of an angle of $\theta$ radians, that $\sin x$ denotes the sine of an angle of $x$ radians, and so on. Likewise for $\cos\theta$, $\tan\theta$, etc. If we want to speak about the sine of an angle of $x$ *degrees*, we shall denote it by $\sin x°$.

In order to become thoroughly familiar with the trigonometric functions and with the use of radian measure, we shall discuss the graphs of the sine, cosine, and tangent. We begin with the sine.

First we must have well in mind the radian measure of angles of 0°, 90°, 180°, 270°, 360°. The corresponding radian measurements are 0, $\pi/2$, $\pi$, $3\pi/2$, $2\pi$. It is also convenient to have in mind the radian equivalents of 30°, 45°, and 60°. All of these are shown in the accompanying table.

| $\phi$ (degrees): | 0 | 30 | 45 | 60 | 90 | 180 | 270 | 360 |
|---|---|---|---|---|---|---|---|---|
| $\theta$ (radians): | 0 | $\frac{\pi}{6}$ | $\frac{\pi}{4}$ | $\frac{\pi}{3}$ | $\frac{\pi}{2}$ | $\pi$ | $\frac{3\pi}{2}$ | $2\pi$ |

Now, as $\theta$ increases from 0 to $\pi/2$, $\sin\theta$ increases from 0 to 1; then, as $\theta$ increases from $\pi/2$ to $\pi$, $\sin\theta$ decreases from 1 to 0. As $\theta$ goes from $\pi$ to

$2\pi$, $\sin\theta$ goes through negative values, from 0 to $-1$ and back to 0, with $\sin 3\pi/2 = -1$. The full range of values of $\sin\theta$ is displayed as $\theta$ goes from 0 to $2\pi$. We get a repetition of the same pattern as $\theta$ goes from $2\pi$ to $4\pi$, from $-2\pi$ to 0, or through any other such interval of length $2\pi$. Because of this the sine function is said to be *periodic*, with the period $2\pi$. The graph is shown in Fig. 4-2.

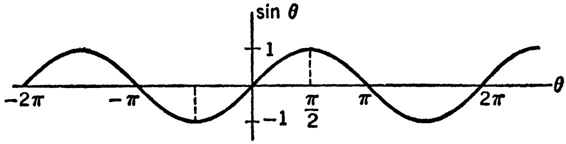



**Fig. 4-2**

The cosine function also has the period $2\pi$. The values of $\cos\theta$ oscillate from $-1$ to $+1$ in the same manner as the values of $\sin\theta$, but $\cos 0 = 1$, $\cos\pi/2 = 0$. The graph of $\cos\theta$ is obtained if the graph of $\sin\theta$ is trans-

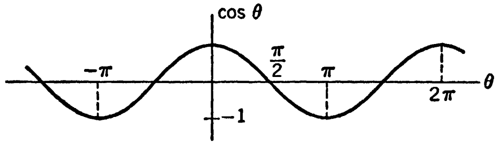



**Fig. 4-3**

lated $\pi/2$ units in the direction of the negative $\theta$-axis (see Fig. 4-3). This relation between the graphs of $\cos\theta$ and $\sin\theta$ is made clear by the fact that

$$\cos\theta = \sin\left(\theta + \frac{\pi}{2}\right).$$

The tangent function is quite different from the sine and cosine in its behavior. The full range of values of $\tan\theta$ is displayed as $\theta$ takes on all

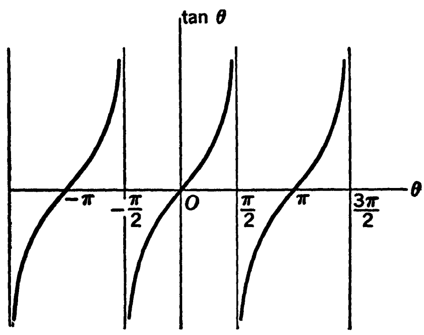



**Fig. 4-4**

values such that $-\pi/2 < \theta < \pi/2$, and there is a repetition of the pattern in the intervals $\pi/2 < \theta < 3\pi/2$, $-3\pi/2 < \theta < -\pi/2$, etc. When $\theta$ is an odd multiple of $\pi/2$, $\tan \theta$ is not defined. As $\theta$ approaches $\pi/2$ from the left side, $\tan \theta \to +\infty$, while $\tan \theta \to -\infty$ as $\theta$ approaches $\pi/2$ from the right side. Each of the lines $\theta = n\pi/2$ ($n$ an odd integer) is a vertical asymptote of the graph, which is shown in Fig. 4-4. The graph exhibits the fact that $\tan \theta$ has the period $\pi$. The analytic statement of this periodicity is given by the formula

$$\tan (\theta + \pi) = \tan \theta,$$

whose validity is evident if we put $A = \theta$, $B = \pi$ in the formula for $\tan (A + B)$.

For convenience we present a brief table of the values of $\sin \theta$, $\cos \theta$, $\tan \theta$. There is no entry for $\tan \theta$ in the $\theta = \pi/2$ column, because $\tan \pi/2$ is undefined.

| $\theta$ | 0 | $\pi/6$ | $\pi/4$ | $\pi/3$ | $\pi/2$ |
|---|---|---|---|---|---|
| $\sin \theta$ | 0 | $1/2$ | $\sqrt{2}/2$ | $\sqrt{3}/2$ | 1 |
| $\cos \theta$ | 1 | $\sqrt{3}/2$ | $\sqrt{2}/2$ | $1/2$ | 0 |
| $\tan \theta$ | 0 | $\sqrt{3}/3$ | 1 | $\sqrt{3}$ | — |

The remaining three trigonometric functions: $\text{ctn}\, \theta$, $\sec \theta$, $\csc \theta$, are most conveniently studied through their definitions in terms of the sine and cosine. See the trigonometric review in § 1-3. Construction of the graphs of these functions is left as an exercise.

The sine and cosine functions are continuous for all values of $\theta$; the other trigonometric functions are continuous for all values of $\theta$ for which they are defined.

## EXERCISES

**1.** (a) Make a careful graph of $\csc \theta$. Begin by considering $0 \le \theta \le \pi$, and work from the graph of $\sin \theta$. Note especially the values of $\csc \theta$ for $\theta = \pi/6, \pi/2, 5\pi/6$. For what values of $\theta$ is $\csc \theta$ *not* defined? Next consider $\pi \le \theta \le 2\pi$ and other intervals between consecutive multiples of $\pi$. Describe features of symmetry and patterns of alternation and periodicity you observe in the graph. Are there any vertical asymptotes?

(b) Make a careful graph of $\sec \theta$. How is the graph related to that of $\csc \theta$?

(c) Make a careful graph of $\text{ctn}\, \theta$. For what values of $\theta$ is $\text{ctn}\, \theta$ not defined? What vertical asymptotes are there? What periodicity does the graph show? Describe a procedure for drawing the graph of ctn $\theta$ by using a tracing of the graph of tan $\theta$. One such procedure is based on the formula

$$\tan\left(\frac{\pi}{2} - \theta\right) = \text{ctn}\ \theta.$$

**2.** In drawing the graphs of the trigonometric functions we tacitly use the continuity of these functions when we draw the curves without breaks except at the asymptotes. The assertion made in the text about continuity of the trigonometric functions can be justified by the following steps, details of which are to be carried out by the student.

(a) The sine function is continuous at $\theta = 0$. This means the same as $\lim_{\theta \to 0} \sin\theta = 0$, because $\sin 0 = 0$. For proof of this limit assertion it will suffice to establish that $|\sin\theta| \le |\theta|$ if $|\theta| < \pi/2$. The student should explain why this inequality will be true for $-\pi/2 < \theta < 0$ if it is true for $0 < \theta < \pi/2$. Then he should use Fig. 4-5 to explain why $0 < \sin\theta < \theta$ if $0 < \theta < \pi/2$.

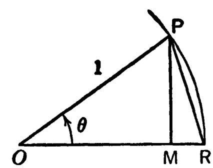



**Fig. 4-5**

(b) The cosine function is continuous at $\theta = 0$. This is the same as the assertion that $\lim_{\theta \to 0} \cos\theta = 1$. Why? For proof use Fig. 4-5 to show that $0 < 1 - \cos\theta < \theta$ if $0 < \theta < \pi/2$. Then explain why $|\cos\theta - 1| \le |\theta|$ if $|\theta| < \pi/2$. This inequality implies the required limit assertion.

(c) The sine function is continuous for all values of $\theta$. For, if $\theta$ is any fixed value, and if we write $\theta - \theta_0 = h$, then

$$\sin\theta = \sin(\theta_0 + h) = \sin\theta_0 \cos h + \cos\theta_0 \sin h.$$

What theorems about limits are now needed, along with the results of (a) and (b), to show that $\lim_{\theta \to \theta_0} \sin\theta = \sin\theta_0$?

(d) Carry out an argument like that in (c) to show that the cosine function is continuous for all values of $\theta$.

(e) What theorem about limits is needed, along with the results in (c) and (d), to prove the continuity of the tangent function and of the remaining trigonometric functions for all values of $\theta$ for which they are defined?

## 4-2 Derivatives of the Sine and Cosine

In this section we shall show how to discover formulas for $\frac{d}{dx}\sin x$ and $\frac{d}{dx}\cos x$. The properties of the sine and cosine functions are such that if we are able to find their derivatives at the point $x = 0$, we can at once find the derivatives at all other points.

Hence we shall begin by trying to find $f'(0)$, where $f(x) = \sin x$. Now $f(0) = \sin 0 = 0$, and so by definition

$$f'(0) = \lim_{\Delta x \to 0} \frac{f(0 + \Delta x) - f(0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\sin \Delta x}{\Delta x}.$$

We shall prove that $f'(0) = 1$. If we use $\theta$ instead of $\Delta x$ as an independent variable, we have to show that

$$\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1. \tag{1}$$

Since the ratio $(\sin \theta)/\theta$ is unaltered if $\theta$ is replaced by $-\theta$, it suffices to prove (1) on the assumption that $\theta > 0$; and since $\theta \to 0$, we can assume $\theta < \pi/2$. We now work from Fig. 4-6, in which the circle has unit radius, so that

$$\sin \theta = MP, \quad \cos \theta = OM, \quad \tan \theta = NR.$$

Now, the area of the sector $NOP$ is $\frac{1}{2}\theta$ (being the fraction $\theta/2\pi$ of the total circle), and this area is clearly larger than that of the triangle $MOP$, but smaller than that of the triangle $NOR$. In other words,

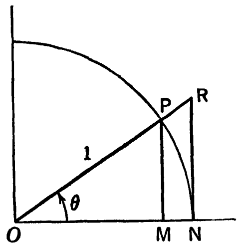


Fig. 4-6

$$\tfrac{1}{2} \sin \theta \cos \theta < \tfrac{1}{2}\theta < \tfrac{1}{2} \tan \theta.$$

If we drop the factor $\frac{1}{2}$ and take the reciprocal quantities, the inequalities go in the reverse order:

$$\frac{1}{\tan \theta} = \frac{\cos \theta}{\sin \theta} < \frac{1}{\theta} < \frac{1}{\sin \theta \cos \theta}.$$

Finally, multiplying through by $\sin \theta$, we obtain

$$\cos \theta < \frac{\sin \theta}{\theta} < \frac{1}{\cos \theta}.$$

Now suppose that $\theta \to 0$. Then $\cos \theta \to 1$ and $1/\cos \theta \to 1$ also. Hence $\sin \theta/\theta$ must approach 1, for it is in between two things, each of which approaches 1. This proves (1).

It is also important for us to know that

$$\lim_{\theta \to 0} \frac{\cos \theta - 1}{\theta} = 0. \tag{2}$$

To prove this we use the half-angle formula

$$1 - \cos \theta = 2 \sin^2 \frac{\theta}{2},$$

from which

$$\frac{\cos \theta - 1}{\theta} = \frac{-2 \sin^2 \frac{\theta}{2}}{\theta} = -\frac{\sin \frac{\theta}{2}}{\frac{\theta}{2}} \cdot \sin \frac{\theta}{2} = -\frac{\sin t}{t} \cdot \sin t,$$

where $t = \theta/2$. If $\theta \to 0$, then $t \to 0$ also, and so $\sin t \to 0$ and $\sin t/t \to 1$. Then, since the limit of the product is the product of the limits, we see that

$$\lim_{\theta \to 0} \frac{\cos \theta - 1}{\theta} = -1 \cdot 0 = 0.$$

Suppose now that $g(x) = \cos x$. To find the derivative at $x = 0$, we have

$$g'(0) = \lim_{\Delta x \to 0} \frac{g(0 + \Delta x) - g(0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\cos \Delta x - 1}{\Delta x}.$$

By (2) we see that $g'(0) = 0$.

Now we shall deduce the general formulas for the derivatives of $\sin x$ and $\cos x$. We start from the addition formulas

$$\sin (x + \Delta x) = \sin x \cos \Delta x + \cos x \sin \Delta x,$$
$$\cos (x + \Delta x) = \cos x \cos \Delta x - \sin x \sin \Delta x.$$

Then we form the difference quotients

$$\frac{\sin (x + \Delta x) - \sin x}{\Delta x} = \sin x \frac{\cos \Delta x - 1}{\Delta x} + \cos x \frac{\sin \Delta x}{\Delta x},$$
$$\frac{\cos (x + \Delta x) - \cos x}{\Delta x} = \cos x \frac{\cos \Delta x - 1}{\Delta x} - \sin x \frac{\sin \Delta x}{\Delta x}.$$

As $\Delta x \to 0$ we can use (1) and (2) with $\Delta x$ in place of $\theta$; the results are

$$\lim_{\Delta x \to 0} \frac{\sin (x + \Delta x) - \sin x}{\Delta x} = \frac{d}{dx} \sin x = \cos x, \tag{3}$$

and

$$\lim_{\Delta x \to 0} \frac{\cos (x + \Delta x) - \cos x}{\Delta x} = \frac{d}{dx} \cos x = -\sin x.$$

When these results are combined with the composite function rule (Theorem 3-E), we have

$$\frac{d}{dx} \sin u = \cos u \frac{du}{dx}, \tag{4}$$

$$\frac{d}{dx} \cos u = -\sin u \frac{du}{dx}, \tag{5}$$

where $u$ denotes any differentiable function of $x$.

**Example 1:** Find $\frac{dy}{dx}$ if $y = \sin 3x^2$. Here $u = 3x^2$, so

$$\frac{dy}{dx} = \cos 3x^2 \frac{d}{dx} (3x^2) = 6x \cos 3x^2.$$

**Example 2:**

$$\frac{d}{dx}\cos\sqrt{1-x} = -\sin\sqrt{1-x}\,\frac{d}{dx}\sqrt{1-x}$$
$$= \frac{1}{2\sqrt{1-x}}\sin\sqrt{1-x}.$$

The student must learn to use formulas (4) and (5) in conjunction with all previous rules of differentiation. We remind the student of the convention of notation for powers of the trigonometric functions: In general $\sin^n x$ means $(\sin x)^n$. There is one exception to this rule, however; $(\sin x)^{-1}$ is *never* written as $\sin^{-1} x$, for this latter expression is regularly used for the *inverse sine* function of $x$.

**Example 3:** Find $\frac{dy}{dx}$ if $y = \sin^3 5x$. Let $w = \sin 5x$. Then $y = w^3$, and

$$\frac{dy}{dx} = 3w^2\frac{dw}{dx} = 3\sin^2 5x \cdot \cos 5x\,\frac{d}{dx}(5x),$$

$$\frac{d}{dx}\sin^3 5x = 3\sin^2 5x \cdot 5\cos 5x = 15\sin^2 5x\cos 5x.$$

We can now explain the reason for preferring radian measure instead of degree measure when dealing with the trigonometric functions in calculus. Let $\sin x^\circ$ and $\cos x^\circ$ denote the sine and cosine of an angle of $x$ degrees. Since an angle of $x$ degrees contains $\pi x/180$ radians,

$$\sin x^\circ = \sin\frac{\pi x}{180}.$$

Then
$$\frac{d}{dx}\sin x^\circ = \cos\frac{\pi x}{180}\,\frac{d}{dx}\left(\frac{\pi x}{180}\right) = \frac{\pi}{180}\cos\frac{\pi x}{180},$$

or
$$\frac{d}{dx}\sin x^\circ = \frac{\pi}{180}\cos x^\circ.$$

This formula takes the place of the simpler formula (3), which holds when we understand $\sin x$ and $\cos x$ to be defined with reference to an angle of *$x$ radians*. It is to avoid the repeated occurrence of the factor $\pi/180$ that we use radian measure.

## EXERCISES

**1.** Find $\frac{dy}{dx}$ in each case.

(a) $y = 2\sin(5x - 7)$.
(b) $y = 5\cos(2x - 3)$.
(c) $y = \sin\sqrt{x}$.
(d) $y = \cos 2(3x - 4)^2$.
(e) $y = \cos^3 x - \sin^3 x$.
(f) $y = x^3\cos 2x - x^2\sin 3x$.
(g) $y = \sin x\cos^2 x$.
(h) $y = \cos^2 2x\sin^3 2x$.
(i) $y = x^2\sin\frac{1}{x}$.
(j) $y = x^4\sin\frac{1}{x^2}$.
(k) $y = (1 - 2\sin^2 3x)^{1/2}$.
(l) $y = (3\cos^2 4x + 1)^{3/2}$.

2. Find $\frac{dy}{dx}$ and $\frac{d^2y}{dx^2}$ in each case.

(a) $y = \sin^5 x$.

(b) $y = \cos^4 2x$.

(c) $y = \frac{\sin x}{1 - \sin x}$.

(d) $y = \frac{1 - \cos x}{1 + \cos x}$.

(e) $y = \frac{27}{\sin x} + \frac{64}{\cos x}$.

(f) $y = 4 \sin x \cos^2 x$.

(g) $y = \frac{1}{5 - 3 \cos 2x}$.

(h) $y = \frac{\sin 2x}{x}$.

(i) $y = \frac{1 - \cos 3x}{\sin 3x}$.

(j) $y = 2(\sin x - x \cos x)$.

3. Quite often in calculus the use of trigonometric identities is quite helpful in simplifying the form of functions or their derivatives. The half- and double-angle formulas are often convenient:

$$\sin 2\theta = 2 \sin \theta \cos \theta. \qquad \cos 2\theta = \cos^2 \theta - \sin^2 \theta.$$

$$\sin^2 \frac{\theta}{2} = \frac{1 - \cos \theta}{2}. \qquad \cos^2 \frac{\theta}{2} = \frac{1 + \cos \theta}{2}.$$

In the following exercises various identities may be needed. Show that each derivative can be put in the form indicated.

(a) $\frac{d}{dx}\left(\frac{\sin x - x \cos x}{\cos x}\right) = \tan^2 x$.

(b) $\frac{d}{dx}\left(x + \frac{\cos x}{\sin x}\right) = -\text{ctn}^2 x$.

(c) $\frac{d}{dx}\left(\frac{x}{2} + \frac{\sin 2ax}{4a}\right) = \cos^2 ax$.

(d) $\frac{d}{dx}\left(\frac{x}{2} - \frac{\sin 2ax}{4a}\right) = \sin^2 ax$.

(e) $\frac{d}{dx}\left(\frac{1}{a} \sin ax - \frac{1}{3a} \sin^3 ax\right) = \cos^3 ax$.

(f) $\frac{d}{dx}\left(\frac{x}{8} - \frac{\sin 4ax}{32a}\right) = \sin^2 ax \cos^2 ax$.

(g) $\frac{d}{dx}(1 + \cos 2x)^2 = -16 \cos^3 x \sin x$.

(h) $\frac{d}{dx}\left(2 \cos ax + \frac{1}{2} \sin 2ax \sin ax\right) = -3a \sin^3 ax$.

4. Draw the graph of each equation. Locate the points of zero slope. Find the absolute maxima and minima of $y$. Show that each point where $y = 0$ is a point of inflection. What is the periodicity of $y$ as a function of $x$?

(a) $y = \sin x + \cos x$.

(b) $y = \sin 2x + \sqrt{3} \cos 2x$.

(c) $y = 2\sqrt{3} \sin (x/2) - 2 \cos (x/2)$.

5. Draw the graph of $y = f(x)$, where $f(x) = x - \sin x$. Show that $f(x)$ always increases as $x$ increases. Note that $f'(x)$ is periodic, although $f(x)$ is not. Discuss points of inflection and the sense of concavity at various parts of the curve.

6. (a) Draw the graph of $y = \sqrt{2}x + 4 \cos (x/2)$ for $0 \leq x \leq 2\pi$. Find the absolute maximum and minimum values of $y$ for the specified values of $x$. Are there any points of inflection?
(b) Proceed as in (a) with $y = x + 2 \cos (x/2)$.
(c) Proceed as in (a) with $y = 2x + 3 \cos (x/2)$.
(d) If $a$ and $b$ are positive, show that the absolute minimum of $y = ax + b \cos (x/2)$ on the interval $0 \leq x \leq \pi$ *always* occurs at one of the ends of the interval, while the absolute maximum *never* occurs at $x = 0$. Show also that the absolute maximum, for $0 \leq x \leq \pi$, occurs *inside* the interval if $2a < b$, and at $x = \pi$ if $2a \geq b$. What can you say about the sense of concavity of the graph, in all cases?

## 4-3 Differentiation of the Other Trigonometric Functions

We can find the derivative of $\tan x$ by using the rule for differentiating a quotient:

$$\tan x = \frac{\sin x}{\cos x}, \quad \frac{d}{dx} \tan x = \frac{\cos x (\cos x) - \sin x (-\sin x)}{\cos^2 x}$$

$$\frac{d}{dx} \tan x = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.$$

We can find the derivative of $\operatorname{ctn} x$ in a similar way. The result is

$$\frac{d}{dx} \operatorname{ctn} x = -\csc^2 x.$$

To deal with $\sec x$ we treat it as the reciprocal of $\cos x$:

$$\frac{d}{dx} \frac{1}{\cos x} = \frac{-1}{\cos^2 x} \frac{d}{dx} (\cos x) = \frac{\sin x}{\cos^2 x}.$$

This is usually written in the form

$$\frac{d}{dx} \sec x = \sec x \tan x.$$

The analogous formula for the derivative of $\csc x$ is

$$\frac{d}{dx} \csc x = -\csc x \operatorname{ctn} x.$$

We now compile the following list, in which $u$ denotes an arbitrary differentiable function of $x$:

I. $\dfrac{d}{dx} \sin u = \cos u \dfrac{du}{dx}$.  II. $\dfrac{d}{dx} \cos u = -\sin u \dfrac{du}{dx}$.

III. $\dfrac{d}{dx} \tan u = \sec^2 u \dfrac{du}{dx}$.  IV. $\dfrac{d}{dx} \operatorname{ctn} u = -\csc^2 u \dfrac{du}{dx}$.

V. $\dfrac{d}{dx}\sec u = \sec u \tan u \dfrac{du}{dx}.$ VI. $\dfrac{d}{dx}\csc u = -\csc u \operatorname{ctn} u \dfrac{du}{dx}.$

It is desirable to have these six formulas thoroughly memorized.

**Example 1:** Find $\dfrac{dy}{dx}$ and $\dfrac{d^2y}{dx^2}$ if $y = \operatorname{ctn}\sqrt{x}$. We have

$$\frac{dy}{dx} = -\csc^2\sqrt{x}\,\frac{d}{dx}\sqrt{x} = -\frac{1}{2\sqrt{x}}\csc^2\sqrt{x},$$

$$\frac{d^2y}{dx^2} = -\frac{1}{2\sqrt{x}}\,2\csc\sqrt{x}\,\frac{d}{dx}\csc\sqrt{x} - \frac{1}{2}\csc^2\sqrt{x}\,\frac{d}{dx}\frac{1}{\sqrt{x}}$$

$$= -\frac{1}{\sqrt{x}}\csc\sqrt{x}\cdot\left(-\csc\sqrt{x}\operatorname{ctn}\sqrt{x}\cdot\frac{1}{2\sqrt{x}}\right) - \frac{1}{2}\csc^2\sqrt{x}\cdot\left(-\frac{1}{2}x^{-3/2}\right)$$

$$= \frac{1}{2x}\csc^2\sqrt{x}\operatorname{ctn}\sqrt{x} + \frac{1}{4x^{3/2}}\csc^2\sqrt{x}$$

$$= \frac{\csc^2\sqrt{x}}{4x^{3/2}}(2\sqrt{x}\operatorname{ctn}\sqrt{x} + 1).$$

**Example 2:** Find the minimum value of

$$y = \frac{1}{12}\tan 2x + \frac{1}{4}\operatorname{ctn} 2x$$

for $0 < x < \pi/4$. Is there any point of inflection? Sketch the graph.

We begin by finding $dy/dx$:

$$\frac{dy}{dx} = \frac{1}{12}(\sec^2 2x)\cdot 2 - \frac{1}{4}(\csc^2 2x)\cdot 2$$

$$= \frac{1}{6}\sec^2 2x - \frac{1}{2}\csc^2 2x. \tag{1}$$

To make use of this result it is better to express it in terms of sines and cosines:

$$\frac{dy}{dx} = \frac{1}{6\cos^2 2x} - \frac{1}{2\sin^2 2x} = \frac{\sin^2 2x - 3\cos^2 2x}{6\sin^2 2x\cos^2 2x}. \tag{2}$$

We see that $dy/dx = 0$ if $\sin^2 2x = 3\cos^2 2x$, or, what is the same, when $\tan^2 2x = 3$. We are considering $0 < x < \pi/4$, or $0 < 2x < \pi/2$, so the slope is zero when $\tan 2x = \sqrt{3}$. This means that $2x = \pi/3$, or $x = \pi/6$. From the expression for $dy/dx$ in (2) we see that the slope is negative if $0 < x < \pi/6$ and positive if $\pi/6 < x < \pi/4$. Also, the value of $y$ is large and positive when $x$ is between 0 and $\pi/4$ and near either of these values. Hence the graph has the general appearance shown in Fig. 4-7. The minimum value of $y$ occurs when $x = \frac{\pi}{6}$. It is $y = \frac{\sqrt{3}}{6}$. To see if there is a point of inflection, we calculate the second derivative, starting from (1):

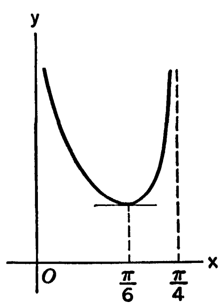



Fig. 4-7

$$\frac{d^2y}{dx^2} = \frac{1}{3}\sec 2x\cdot(\sec 2x \tan 2x)\cdot 2 - \csc 2x\cdot(-\csc 2x \operatorname{ctn} 2x)\cdot 2$$
$$= \frac{2}{3}\sec^2 2x \tan 2x + 2\csc^2 2x \operatorname{ctn} 2x.$$

Since this is always positive for the values of $x$ which we are considering, the curve is concave upward and there is no point of inflection.

## EXERCISES

1. Find $y'$ and $y''$ in each case.

(a) $y = \tan^3 2x.$

(b) $y = \tan x^3.$

(c) $y = \sec^2 5x.$

(d) $y = \operatorname{ctn}\frac{8}{x}.$

(e) $y = \frac{\csc 3x}{x}.$

(f) $y = x \operatorname{ctn}^2 2x.$

(g) $y = \sqrt{1 + \tan 3x}.$

(h) $y = \operatorname{ctn}\frac{x}{x-1}.$

2. In the right triangle shown in Fig. 4-8 suppose that $\theta$ is decreasing at the rate of $\frac{1}{30}$ radian per second. Find each of the indicated derivatives if additional conditions are as specified.

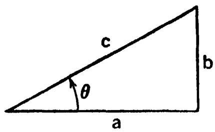


Fig. 4-8

(a) Find $\frac{db}{dt}$ when $\theta = \frac{\pi}{3}$ if $a$ is constantly 12 inches.

(b) Find $\frac{dc}{dt}$ when $\theta = \frac{\pi}{4}$ if $b$ is constantly $10\sqrt{2}$ inches.

(c) Find $\frac{da}{dt}$ when $b = 20$ feet, if $c$ is constantly 40 feet.

(d) Find $\frac{db}{dt}$ when $b = 20$ feet, if $c$ is constantly 40 feet.

(e) Find $\frac{dc}{dt}$ when $b = a$ if $a$ remains 1 mile at all times.

(f) Find $\frac{da}{dt}$ when $a = 1$ foot and $c = 2$ feet if $a$ and $b$ are both changing and $b$ is increasing at the rate of $\frac{2}{15}$ foot per second.

3. A point $P$ is tracing out a circle of radius 10 feet with center at the origin, at the rate of 1 revolution per minute (see Fig. 4-9). The tangent at $P$ intersects the $x$-axis at $Q$. How fast is $Q$ moving when it is 20 feet from $O$?

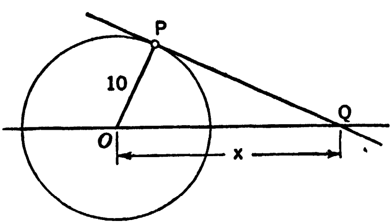


Fig. 4-9

4. A bomber $B$ is flying 400 miles per hour on a level course 2 miles above level land (see Fig. 4-10). The bombardier is sighting on a target $C$. At the instant when the angle of depression between the plane's path and the bombardier's line of sight is 30°, how fast must this line of sight be turning in order to keep on the target? Obtain the answer at first in radians per hour, and then convert to degrees per second.

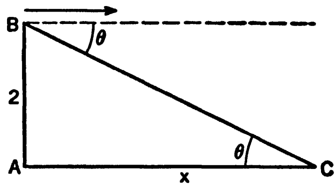



**Fig. 4-10**

5. A lighthouse has a revolving light which turns at the rate of 2 revolutions per minute. The lighthouse is situated $\frac{1}{2}$ mile from a straight beach. Find how fast the spot of light from the beam is moving along the beach when it is 1 mile from the point of the beach nearest the light.

6. A ladder 10 feet long leans against a house. The upper end slips down the wall 5 feet per second. How fast is the ladder turning when it makes an angle of 30° with the ground?

7. A ferris wheel 50 feet in diameter makes 1 revolution every 2 minutes. If the center of the wheel is 30 feet above the ground, how fast is a passenger in the wheel moving vertically when he is $42\frac{1}{2}$ feet above the ground? How fast is he moving horizontally at the same moment?

8. Sketch the graph of $y = \tan x/(1 + \tan x)$ for $0 \le x < \pi/2$. What happens to $y$ and $y'$ as $x \to \pi/2$? Is there a point of inflection? Discuss the sense of concavity and show it on the graph.

9. Sketch the graph of $y = 10 \csc x - 5 \operatorname{ctn} x$ when $0 < x \le \pi/2$, and find the value of $x$ for which $y$ is least. Is there a point of inflection?

10. Sketch the graph of $y = 3 \sec x + 4 \cos x$ when $0 \le x < \pi/2$, and find the minimum value of $y$. Show that there is a point of inflection when $\sin x = 1/\sqrt{7}$.

## 4-4 The Inverse Trigonometric Functions

In working problems in trigonometry we sometimes have occasion to raise and answer questions of the following sort:

**Example 1:** What angle has its sine equal to 0.2419?

If we are studying triangles when we ask this question, it is understood that the angle in question must be between 0° and 180°, or in radian measure, between 0 and $\pi$. But we may be studying "general angle" trigonometry, and then this question may be answered by giving *any* appropriate angle, either positive or negative and without restriction of size. Thus, if $\sin \theta = 0.2419$, we know that $\theta$ is either a first or second quadrant angle, because the sine is positive. From Table IV at the back of this book we see that $\theta = 0.2443$ radians is one possibility. To this may be added or subtracted any integral multiple of $2\pi$, and in this way we get all the first quadrant angles whose sines are 0.2419. Another solution of the problem is $\theta = \pi - 0.2419$, because $\sin(\pi - \theta) = \sin \theta$. Here again we may add or subtract any integral multiple of $2\pi$.

It must be realized, of course, that when we deal with numerical tables we are usually dealing in approximations. The four-decimal-place values of the sines in Table IV are merely approximations to the sines of the indicated angles. Or, if we regard the four-digit decimal sine as an exact thing, then the angle corresponding to it in the table is given only approximately.

**Example 2:** What is $\theta$ if $\tan \theta = -\sqrt{3}$?

Here we know that $\theta$ must be a second or fourth quadrant angle, because the tangent is negative. Now $-\sqrt{3}$ happens to indicate a special angle, as we see from Fig. 4-11. One solution is $\theta = \pi - (\pi/3) = 2\pi/3$. Another solution is the fourth quadrant angle differing from $2\pi/3$ by $\pi$ radians: $\theta = 5\pi/3$ (or, alternatively, $\theta = -\pi/3$). Since the tangent function has the period $\pi$, *all* values of $\theta$ such that $\tan \theta = -\sqrt{3}$ are given by $\theta = (2\pi/3) + n\pi$, where $n$ can be assigned values $0, \pm 1, \pm 2, \cdots$.

**Example 3:** What is $\theta$ if $\cos \theta = -\frac{1}{2}$?

In this case we know that $\theta$ is in either the second or third quadrant. One second quadrant possibility is $\theta = 2\pi/3$, as we see from Fig. 4-11. Since $\cos(-\theta) = \cos \theta$, a third quadrant possibility is $\theta = -2\pi/3$. The complete answer to our problem is then displayed in the form

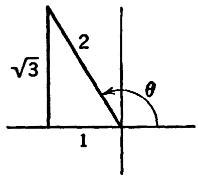



**Fig. 4-11**

$$\theta = \pm\frac{2\pi}{3} + 2\pi n, \qquad n = 0, \pm 1, \pm 2, \cdots.$$

Experience has shown that it is convenient to make a standardized agreement upon what may be called the *principal angle* whose sine is a given number, and likewise for the cosine, tangent, and cotangent. These agreements about principal angles are bound up with the definitions of the inverse trigonometric functions. We take up these definitions one at a time.

*The Inverse Sine Function*

Any number $x$ such that $|x| \leq 1$ can be regarded as the sine of many different angles. By the *principal angle* $y$ whose sine is $x$ we shall always mean that $y$ for which $-\pi/2 \leq y \leq \pi/2$ and $\sin y = x$. We then call $y$

the *inverse sine* of $x$, and write $y = \sin^{-1} x$. The reason for choosing $y$ between $-\pi/2$ and $\pi/2$ is made clear if we examine the graph of the sine function. It is certainly natural and convenient to choose $y$ between 0 and $\pi/2$ if $0 \leq x \leq 1$. Once this choice is made, the choice of $y$ corresponding to an $x$ between $-1$ and 0 *must* be made with $y$ between $-\pi/2$ and 0 if $y = \sin^{-1} x$ is to yield a function which is continuous when $-1 \leq x \leq 1$. The graph of $x = \sin y$, with $y$ as the independent variable, is shown in Fig. 4-12. The part for which $|y| \leq \pi/2$ is fully drawn, and the

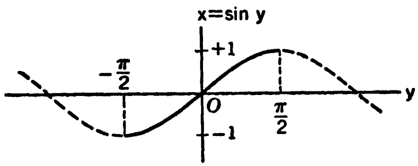



Fig. 4-12

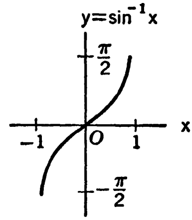



Fig. 4-13

rest of the graph is indicated in dashed form. If we now alter our point of view, and regard $x$ as independent and $y$ as dependent, we obtain the graph of $y = \sin^{-1} x$ (see Fig. 4-13). This graph has free ends at the points $(1, \pi/2)$ and $(-1, -\pi/2)$.

In elementary books on trigonometry, and sometimes in books on other subjects, the inverse sine is regarded as a many-valued function. In such cases, when some agreement is made to select a particular value from among the many, the selected value is called the principal value. In such usage, what we have called simply the inverse sine is called the principal value of the inverse sine. For us in this book the inverse sine will always be, as we have defined it, a single-valued function.

The notation arc sin $x$ is rather widely used as an alternative for $\sin^{-1} x$. The $-1$ in the notation $\sin^{-1} x$ is *not* an exponent. It is merely part of the conventional notation for the inverse sine function.

### *The Inverse Cosine Function*

In defining the inverse cosine we follow the same principle that governed our definition of the inverse sine. We select an unbroken portion

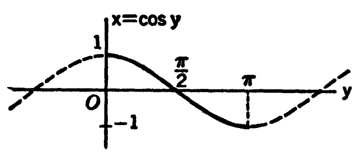



Fig. 4-14

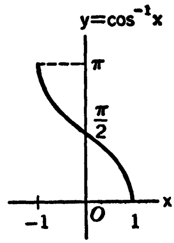



Fig. 4-15

of the graph of $x = \cos y$, in such a way that all values of the cosine are represented once and only once. The most convenient such portion is that for which $0 \leq y \leq \pi$, and it is this portion that we select (see Fig. 4-14). We define $y = \cos^{-1} x$ to mean that $x = \cos y$ and $0 \leq y \leq \pi$. The graph of the inverse cosine is shown in Fig. 4-15.

*The Inverse Tangent Function*

Every real number is a value of the tangent function, and we obtain each value exactly once if we restrict the independent variable to values

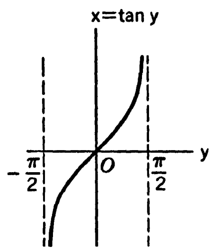



**Fig. 4-16**

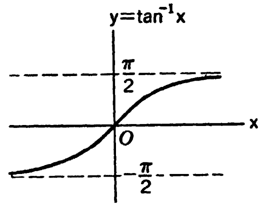



**Fig. 4-17**

between $-\pi/2$ and $\pi/2$ (see Fig. 4-16). Therefore we define $y = \tan^{-1} x$ to mean that $x = \tan y$ and $-\pi/2 < y < \pi/2$. The graph of the inverse tangent is shown in Fig. 4-17.

*The Inverse Cotangent Function*

We define $y = \text{ctn}^{-1} x$ to mean that $x = \text{ctn}\, y$ and $0 < y < \pi$. See Fig. 4-18 and Fig. 4-19.

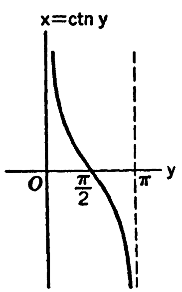



**Fig. 4-18**

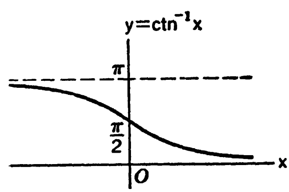



**Fig. 4-19**

The inverse secant and the inverse cosecant are seldom used, and we shall suffer no inconvenience by avoiding the use of them. There is some variation of usage in the definitions of these functions in textbooks. We define the inverse secant and the inverse cosecant, respectively, by the formulas

$$\sec^{-1} x = \cos^{-1}\left(\frac{1}{x}\right), \qquad \csc^{-1} x = \sin^{-1}\left(\frac{1}{x}\right). \tag{1}$$

Thus, in accordance with our previous definitions, the values of $\sec^{-1} x$ lie between 0 and $\pi$, while those of $\csc^{-1} x$ lie between $-\pi/2$ and $\pi/2$. In some books these functions are defined in such a way that their values lie between $-\pi$ and $-\pi/2$ when $x \leq -1$.

**Example 4:**

$$\sin^{-1}\left(\frac{\sqrt{3}}{2}\right) - \sin^{-1}\left(-\frac{1}{2}\right) = \frac{\pi}{3} - \left(-\frac{\pi}{6}\right) = \frac{\pi}{2}.$$

**Example 5:**

$$\cos^{-1}(-1) - \cos^{-1}(1) = \pi - 0 = \pi.$$

**Example 6:**

$$\tan^{-1}(1) - \tan^{-1}(-\sqrt{3}) = \frac{\pi}{4} - \left(-\frac{\pi}{3}\right) = \frac{7\pi}{12}.$$

**Example 7:** Find $\sin^{-1}(-0.6018)$. From Fig. 4-13 we see that

$$\sin^{-1}(-x) = -\sin^{-1} x. \tag{2}$$

Hence, using Table IV, we see that

$$\sin^{-1}(-0.6018) = -\sin^{-1}(0.6018) = -0.6458.$$

We also note, from Fig. 4-17, that

$$\tan^{-1}(-x) = -\tan^{-1} x. \tag{3}$$

**Example 8:** Find $\text{ctn}^{-1}(-0.8050)$. From Fig. 4-19 we see that if $x < 0$, then $\text{ctn}^{-1} x$ exceeds $\pi/2$ by as much as $\pi/2$ exceeds $\text{ctn}^{-1}(-x)$. In other words,

$$\text{ctn}^{-1} x - \frac{\pi}{2} = \frac{\pi}{2} - \text{ctn}^{-1}(-x),$$

or

$$\text{ctn}^{-1} x + \text{ctn}^{-1}(-x) = \pi. \tag{4}$$

In this case, therefore,

$$\text{ctn}^{-1}(-0.8050) = \pi - \text{ctn}^{-1} 0.8050 = \pi - 0.8930 = 2.2486.$$

*Differentiation Formulas*

To find the derivative of the inverse sine function we start with $y = \sin^{-1} x$, so that $-\pi/2 \leq y \leq \pi/2$ and $x = \sin y$. Differentiating this last formula with respect to $x$, we see that

$$1 = \cos y \frac{dy}{dx}, \qquad \frac{dy}{dx} = \frac{1}{\cos y}.$$

Now $\cos^2 y = 1 - \sin^2 y = 1 - x^2$. We know that $\cos y \geq 0$, because of the inequalities imposed on $y$. Therefore $\cos y$ is the *positive* square root of $1 - x^2$. Thus we obtain the formula

$$\frac{d}{dx} \sin^{-1} x = \frac{1}{\sqrt{1 - x^2}}. \tag{5}$$

The derivatives of the other inverse trigonometric functions may be calculated by similar methods. If $y = \cos^{-1} x$, we have $0 \leq y \leq \pi$ and $x = \cos y$. Therefore

$$1 = -\sin y \frac{dy}{dx}, \qquad \frac{dy}{dx} = \frac{-1}{\sin y}.$$

But $\sin y = \sqrt{1 - \cos^2 y} = \sqrt{1 - x^2}$ (the positive square root because $\sin y \geq 0$ for the values of $y$ under consideration). Therefore

$$\frac{d}{dx} \cos^{-1} x = \frac{-1}{\sqrt{1 - x^2}}. \tag{6}$$

For the inverse tangent and inverse cotangent we list the formulas, leaving the derivation as exercises for the student. The formulas are

$$\frac{d}{dx} \tan^{-1} x = \frac{1}{1 + x^2}, \tag{7}$$

$$\frac{d}{dx} \operatorname{ctn}^{-1} x = \frac{-1}{1 + x^2}. \tag{8}$$

In applying these formulas we frequently want to combine them with the use of the chain rule. Hence we make the following formal list, in which $u$ denotes any differentiable function of $x$.

I. $\dfrac{d}{dx} \sin^{-1} u = \dfrac{1}{\sqrt{1 - u^2}} \dfrac{du}{dx}.$

II. $\dfrac{d}{dx} \cos^{-1} u = \dfrac{-1}{\sqrt{1 - u^2}} \dfrac{du}{dx}.$

III. $\dfrac{d}{dx} \tan^{-1} u = \dfrac{1}{1 + u^2} \dfrac{du}{dx}.$

IV. $\dfrac{d}{dx} \operatorname{ctn}^{-1} u = \dfrac{-1}{1 + u^2} \dfrac{du}{dx}.$

**Example 9:**

$$\frac{d}{dx} \sin^{-1} \frac{x}{3} = \frac{1}{\left(1 - \dfrac{x^2}{9}\right)^{1/2}} \cdot \frac{1}{3} = \frac{1}{\sqrt{9 - x^2}}.$$

**Example 10:**

$$\frac{d}{dx} \tan^{-1} \frac{4x - 3}{\sqrt{2}} = \frac{1}{1 + \left(\dfrac{4x - 3}{\sqrt{2}}\right)^2} \cdot \frac{4}{\sqrt{2}} = \frac{4\sqrt{2}}{16x^2 - 24x + 11}.$$

**Example 11:**

$$\frac{d}{dx} \cos^{-1} \frac{1}{x^2} = \frac{-1}{\left(1 - \dfrac{1}{x^4}\right)^{1/2}} \cdot \frac{-2}{x^3} = \frac{2}{x\sqrt{x^4 - 1}}.$$

## EXERCISES

**1.** Give the numerical values of

(a) $\sin^{-1}\left(\frac{-1}{\sqrt{2}}\right) - \sin^{-1}(-1).$

(b) $\text{ctn}^{-1}\left(\frac{\sqrt{3}}{3}\right) - \tan^{-1}(-\sqrt{3}).$

(c) $\cos^{-1}(0) - \sin^{-1}\left(\frac{1}{2}\right).$

(d) $\tan^{-1}(-1) + \cos^{-1}\left(\frac{-1}{2}\right).$

(e) $\cos^{-1}\left(\frac{-\sqrt{3}}{2}\right) - \text{ctn}^{-1}(-1).$

**2.** Use Table IV to evaluate

(a) $\sin^{-1}(-0.6450).$

(b) $\cos^{-1}(-0.3007).$

(c) $\tan^{-1}(-2.211).$

(d) $\text{ctn}^{-1}(-4.011).$

**3.** Explain the relation between $\cos^{-1}(-x)$ and $\cos^{-1} x$. Use the relation to find $\cos^{-1}(-0.5878)$.

**4.** Find $f'(x)$ if

(a) $f(x) = \sin^{-1} x^2.$

(b) $f(x) = \cos^{-1} 3x.$

(c) $f(x) = \tan^{-1}\frac{x}{5}.$

(d) $f(x) = \text{ctn}^{-1}\frac{1}{x}.$

**5.** Find $y'$ if

(a) $y = \sin^{-1}\sqrt{x}.$

(b) $y = \cos^{-1}\frac{x-2}{2}.$

(c) $y = \cos^{-1}\frac{8x-9}{9}.$

(d) $y = \text{ctn}^{-1}\frac{2x}{1-x^2}.$

(e) $y = \tan^{-1}\frac{x-1}{x+1}.$

(f) $y = \sin^{-1}\frac{2x+1}{\sqrt{3}}.$

(g) $y = \text{ctn}^{-1}\frac{1-x}{1+x}.$

(h) $y = \tan^{-1}\frac{6x}{9-x^2}.$

**6.** (a) If $y = \tan\left(2\tan^{-1}\frac{x}{2}\right)$, show that

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

(b) If $y = x\tan^{-1}\frac{x}{y}$, show that

$$\frac{dy}{dx} = \frac{y}{x}.$$

**7.** Find the following limits:

(a) $\lim_{x \to -\infty} \dfrac{\tan^{-1} x}{\operatorname{ctn}^{-1} x}$; (b) $\lim_{x \to +\infty} (\tan^{-1} x)(1 + \operatorname{ctn}^{-1} x)$;

(c) the limits of $\tan^{-1}\left(2 \tan \dfrac{x}{2}\right)$ as $x \to \pi$ from the left and right, respectively.

**8.** In a right triangle $ABC$ with right angle at $C$, side $AC$ is $b$ units long and side $BC$ is $2x$ units long. If $P$ is the mid-point of $BC$ and $\theta$ is the angle $BAP$, express $\theta$ as a function of $x$. Make a graph of $\theta$ as a function of $x$. What happens to $\theta$ as $x$ gets very large? Find the value of $x$ which makes $\theta$ greatest.

**9.** Points $A$ and $B$ are at $(0, a)$ and $(0, b)$, with $0 < a < b$. Point $P$ is at $(x, 0)$, with $x > 0$, and $\theta$ is the angle $APB$. Express $\theta$ as a function of $x$, using inverse cotangents. Study the behavior of $\theta$ as $x$ goes from very small to very large values, and make a graph. For what value of $x$ is $\theta$ largest?

**10.** Show that

$$\frac{d}{dx} \sin^{-1} \frac{1}{x} = \frac{-1}{|x|\sqrt{x^2 - 1}}, \qquad \frac{d}{dx} \cos^{-1} \frac{1}{x} = \frac{1}{|x|\sqrt{x^2 - 1}}.$$

**11.** Find

(a) $\dfrac{d}{dx} \cos^{-1} \dfrac{1 - x^2}{1 + x^2}$; (b) $\dfrac{d}{dx} \sin^{-1} \sqrt{1 - x^2}$.

**12.** (a) Does the curve $y = x \sin^{-1} x$ have any point of inflection? Does $y$ have any maxima or minima? Sketch the graph.
(b) Proceed as in (a) with $y = x \tan^{-1} x$. Show that the curve has certain lines $y = \pm mx + b$ as asymptotes. What are the correct values of $m$ and $b$?

**13.** Find the point of inflection of the curve $y = (1 + x) \tan^{-1} x$.

**14.** Find the acute angle between the tangents to the curves $y = \tan^{-1} x$, $y = \operatorname{ctn}^{-1} x$ at their point of intersection.

**15.** In each of the following cases express both $y$ and $y'$ in a purely algebraic form. Compute $y'$ in two ways and reconcile the results.
(a) $y = \sin(\cos^{-1} x)$. (c) $y = \sin(\tan^{-1} x)$.
(b) $y = \tan(\cos^{-1} x)$. (d) $y = \operatorname{ctn}(\tan^{-1} x)$.

**16.** Prove that $\tan^{-1} x + \operatorname{ctn}^{-1} x = \pi/2$ by two methods:
(a) Let $y = \operatorname{ctn}^{-1} x$. Show that $-\pi/2 < (\pi/2) - y < \pi/2$ and also that $\tan\left(\dfrac{\pi}{2} - y\right) = x$. What is the conclusion?
(b) Show that $\tan^{-1} x + \operatorname{ctn}^{-1} x$ is constant by considering its derivative and using the proposition V in § 2-1. Then put $x = 0$ to find the value of the constant.

**17.** Using arguments like that in Exercise 16, show that $\sin^{-1} x + \cos^{-1} x = \pi/2$.

**18.** Let $y = \cos^{-1}(\sin x)$. Show that $y' = -1$ if $0 < x < \pi/2$ or $3\pi/2 < x < 2\pi$, while $y' = +1$ if $\pi/2 < x < 3\pi/2$. Then, observing that $y$ is a continuous function of $x$, plot the points of the graph corresponding to $x = 0$, $\pi/2$, $3\pi/2$, $2\pi$, and draw the graph when $0 \le x \le 2\pi$. The derivative is undefined when $x$ is an odd multiple of $\pi/2$.

## 4-5 Maxima and Minima. Rates

In this section we consider a variety of problems in which it is natural to employ trigonometric or inverse trigonometric functions. In principle these problems are like those considered in the latter part of Chapter III (§ 3-10 through § 3-12).

**Example 1:** A vacant plot of ground is situated at the corner of two streets which intersect at right angles. A tree stands at $T$ in the plot, $a$ feet from one street and $b$ feet from the other (see Fig. 4-20). The corner of the

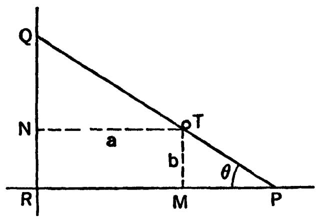



**Fig. 4-20**

plot is to be cut off by a straight fence $PQ$ passing next to the tree. Find the smallest possible value of the area $PRQ$, and the value of $\theta$ which yields this minimum area.

*Solution.* The first step is to express the area $A$ of $PRQ$ as a function of $\theta$. To do this we express $MP$ and $NQ$ in terms of $\theta$:

$$MP = b \operatorname{ctn} \theta, \qquad NQ = a \tan \theta.$$

Then $$A = \frac{1}{2}(a + b \operatorname{ctn} \theta)(b + a \tan \theta).$$

On expanding this, we find

$$A = \frac{1}{2}(2ab + b^2 \operatorname{ctn} \theta + a^2 \tan \theta). \tag{1}$$

By examining the geometry of the situation we can see that $A$ becomes very large when $\theta$ is near 0 and also when $\theta$ is near $\pi/2$. Therefore it is reasonable to expect a minimum value of $A$ for some intermediate value of $\theta$. We differentiate:

$$\frac{dA}{d\theta} = \frac{1}{2}(-b^2 \csc^2 \theta + a^2 \sec^2 \theta) = \frac{a^2 \sin^2 \theta - b^2 \cos^2 \theta}{2 \sin^2 \theta \cos^2 \theta}.$$

From this it is evident that $dA/d\theta = 0$ if $\tan\theta = b/a$, and that the derivative changes from negative to positive as $\theta$ increases through the value $\tan^{-1}(b/a)$. Hence $A$ is smallest for this value of $\theta$. Observe, for this $\theta$, that $M$ and $N$ bisect $RP$ and $RQ$, respectively. To find the actual minimum value of $A$ we put $\tan\theta = b/a$ in (1). The result is:

$$\text{minimum } A = 2ab.$$

**Example 2:** Study the graph of $y = 2 + (1 + \sin x)\cos x$. Find the points of relative maxima and minima, and the points of inflection.

We see that $y$ is a periodic function of $x$, with period $2\pi$. Hence we confine attention to the interval for which $0 \le x \le 2\pi$. The first derivative is

$$\frac{dy}{dx} = (1 + \sin x)(-\sin x) + \cos^2 x.$$

It is convenient to replace $\cos^2 x$ by $1 - \sin^2 x$, so that

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

This makes it easier to tell when the derivative is zero. We can factor; the result is

$$\frac{dy}{dx} = (1 - 2\sin x)(1 + \sin x).$$

This shows that $y' = 0$ if $\sin x = \frac{1}{2}$ (i.e., $x = \pi/6$ or $5\pi/6$) and if $\sin x = -1$ (i.e., $x = 3\pi/2$). Since $1 + \sin x \ge 0$, $y'$ changes sign at $x = \pi/6$ and $x = 5\pi/6$, but not at $x = 3\pi/2$. The sign scheme for $y'$ is:

$$y' \ge 0 \quad \text{if} \quad 0 \le x \le \frac{\pi}{6} \quad \text{or} \quad \frac{5\pi}{6} \le x \le 2\pi,$$

$$y' < 0 \quad \text{if} \quad \frac{\pi}{6} < x < \frac{5\pi}{6}.$$

Hence there is a relative maximum at $x = \pi/6$ and a relative minimum at $x = 5\pi/6$. To draw the graph it suffices to plot the points indicated in the accompanying table. The graph appears in Fig. 4-21.

| $x$ | 0 | $\pi/6$ | $\pi/2$ | $5\pi/6$ | $\pi$ | $3\pi/2$ | $2\pi$ |
|---|---|---|---|---|---|---|---|
| $y$ | 3 | 3.30 | 2 | 0.70 | 1 | 2 | 3 |

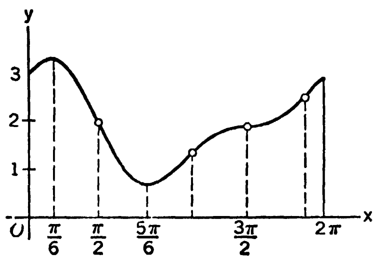



**Fig. 4-21**

To locate points of inflection we have

$$y'' = -\cos x - 4\sin x\cos x = (-\cos x)(1 + 4\sin x).$$

Thus $y''$ changes sign when $\cos x = 0$ (i.e., $x = \pi/2, 3\pi/2$) and when $\sin x = -\frac{1}{4}$ ($x = 3.40$ and $x = 6.03$, approximately). The four points of inflection in the interval are marked by dots on the graph.

It is not always so easy to tell when $y' = 0$ or $y'' = 0$ when $y$ is a trigonometric function of $x$. In general the solution of a trigonometric equation may involve methods of approximation which we are not now ready to study. See §§ 16-2, 16-3, for instance.

**Example 3:** A high tower stands at the end of a level road. A man drives toward the tower at the rate of 60 miles per hour (88 feet per second). The tower rises 500 feet above the level of the man's eyes. How fast is the angle subtended by the tower at the man's eye increasing when the man is 1200 feet from the base of the tower?

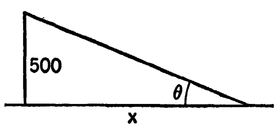


**Fig. 4-22**

*Solution.* The situation is shown in Fig. 4-22. We wish to know the value of $d\theta/dt$ when $x = 1200$, given that $dx/dt = -88$. Now

$$\theta = \operatorname{ctn}^{-1}\frac{x}{500}, \qquad \frac{d\theta}{dt} = \frac{-1}{1+\left(\dfrac{x}{500}\right)^2}\cdot\frac{1}{500}\frac{dx}{dt},$$

$$\frac{d\theta}{dt} = \frac{-500}{(500)^2 + x^2}\frac{dx}{dt}.$$

Substituting, we find

$$\frac{d\theta}{dt} = \frac{500(88)}{(1300)^2} = \frac{440}{16{,}900} = \frac{22}{845}.$$

The units of the answer are radians per second. To convert to degrees per second, multiply by $180/\pi$. The result is approximately 1.5 degrees per second.

As an alternative method of solution, we can write

$$x = 500\operatorname{ctn}\theta, \qquad \frac{dx}{dt} = -88 = -500\csc^2\theta\,\frac{d\theta}{dt},$$

$$\frac{d\theta}{dt} = \frac{88}{500\csc^2\theta} = \frac{88}{500}\sin^2\theta.$$

When $x = 1200$, $\sin\theta = 500/1300 = 5/13$. On substituting, we get the same answer as before.

## EXERCISES

**1.** A variable right circular cylinder is fitted inside a fixed sphere of radius $b$ (see Fig. 4-23). (a) Using the angle $\theta$ as independent variable, express the lateral area of the cylinder as a function of $\theta$, and determine the maximum value of this area. (b) Solve the corresponding problem for the volume of the cylinder.

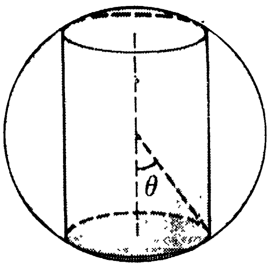



**Fig. 4-23**

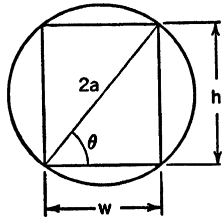



**Fig. 4-24**

**2.** The strength of a rectangular beam cut from a log is proportional to the width $w$ and to the square of the depth $h$. Find the ratio $h/w$ for the strongest beam. See Fig. 4-24.

**3.** Solve Exercise 2 if the strength of the beam is proportional to $wh^3$.

**4.** The longer of the two parallel sides of a trapezoid makes equal angles $\theta$ with the sides adjacent to it. The shorter of the two parallel sides and the other two sides are all of length $b$. What is the maximum possible area of the trapezoid, as $\theta$ varies?

**5.** A right circular cone is inscribed in a sphere of radius 9 inches. (a) Using a figure somewhat analogous to Fig. 4-23, express the volume $V$ of the cone as a function of the angle $\theta$, and find the value of $\theta$ which yields the maximum volume. (b) What is the maximum volume? (c) Draw the graph of $V$ as a function of $\theta$ for $0 \leq \theta \leq \pi$. To what extremes do $\theta = 0$, $\theta = \pi$ correspond?

**6.** Two towns $A$ and $B$ are 8 miles apart. A third town $C$ is located 5 miles from both $A$ and $B$. If the point $P$, equidistant from $A$ and $B$, is such that the sum of the distances $PA$, $PB$, $PC$ is the least possible, how far is it from $C$? Use the angle $ABP$ as independent variable.

**7.** A man is in a boat 1 mile due south of an east-west shore line. Along the shore, $2\frac{1}{2}$ miles east of the point nearest him, is the man's home. The man aims for a point $\theta$ radians east of north, and rows to the shore at the rate of 1.8 miles per hour. He then walks home along the beach at the rate of 3 miles per hour. Express the time $T$ it takes him to reach home as a function of $\theta$, and graph the function. What is the least time in which he can get home, by varying $\theta$?

**8.** (a) A ladder 27 feet long is placed straight up against a fence 8 feet high. The lower end of the ladder is then pulled directly away from the fence. If the ladder is kept in contact with the top of the fence, what is the greatest horizontal distance the ladder ever projects beyond the fence? As independent variable choose the angle which the ladder makes with the ground. (b) Solve the same problem if 27 and 8 are replaced by $c$ and $a$, respectively.

**9.** Find the shortest possible length of the line $PQ$ in Fig. 4-20, which is associated with illustrative Example 1 in the text.

**10.** In Fig. 4-25 is shown a situation where a ray of light travels from $A$ to $P$ with velocity $v_1$ and then, entering a different medium, travels from $P$ to $B$ with velocity $v_2$. The total time $T$ for the ray to travel from $A$ to $B$ is then

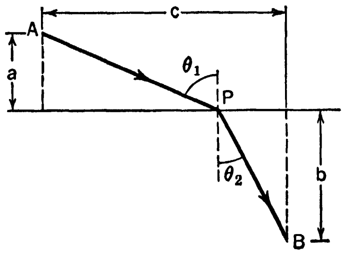



**Fig. 4-25**

dependent on the position of $P$. Express $T$ in terms of $a$, $b$, $v_1$, $v_2$, $\theta_1$, $\theta_2$. Note that $\theta_1$ and $\theta_2$ are connected by the relation $c = a \tan \theta_1 + b \tan \theta_2$. Hence, to make $T$ a minimum by varying the position of $P$ is an extremal problem with a side condition on the two variables $\theta_1$, $\theta_2$, of the general type considered in § 3-11. Choose $\theta_1$ as the independent variable and deduce that when $T$ is a minimum, $\sin \theta_1 / \sin \theta_2 = v_1/v_2$. This equation expresses the optical principle of refraction known as *Snell's law*. The physical law that light travels from one point to another along the path requiring the least time is known as Fermat's principle. This principle applies, not only to obtain Snell's law of refraction, but also to the determination of the paths of light rays in media of variable density, where in general the light will travel along curves, rather than in straight lines.

**11.** Draw the graph of each of the following equations. Confine attention in each case to an interval which displays one complete period of the function of $x$. Locate the relative maxima and minima.

(a) $y = \sin x + \sin x \cos x$.
(b) $y = \sin 2x - \sin 2x \cos 2x$.
(c) $y = 4 \cos x + \cos^2 x - \sin^2 x$.
(d) $y = 4 \cos x + 4 \sin^2 x - 2$.
(e) $y = 1 + 8 \cos x - 2 \cos^2 x$.
(f) $y = 5 \cos^3 x - 3 \cos x$.
(g) $y = 2 \sin 3x + 2 \cos^2 3x - 1$.

**12.** Graph each equation, $0 \leq x \leq 2\pi$, finding all relative extrema and all points of inflection.

(a) $y = 4 \sin x - 4 \sin^3 x$. (b) $y = \dfrac{\cos x}{3 - \sqrt{2} \cos x}$.

**13.** A signboard 45 feet high stands at the top of a cliff 86 feet high. How far from the foot of the cliff should a man stand in order to have the sign subtend the largest possible angle at his eyes, which are 6 feet above the ground?

**14.** A rocket aimed straight up has risen $\frac{1}{12}t^{3/2}$ miles above the earth $t$ seconds after it starts. An observer 4 miles from the launching site is observing the rocket through a telescope. How fast is the angle of elevation of the telescope increasing when $t = 16$ seconds?

**15.** (a) An isosceles triangle has base of length $b$ and equal sides of length $c$. Express the angle at the apex as a function of $b$ and $c$. (b) Find the rate of change of the apex angle when the base is 48 inches and increasing 12 inches per minute, and the sides are 26 inches and increasing 13 inches per minute. (c) Find the rate of change of the base at an instant when the triangle happens to be equilateral, 26 inches on a side, the apex angle is increasing $\sqrt{3}$ radians per minute, and the sides adjacent to the apex are decreasing 19 feet per minute.

**16.** A weight is drawn along a level floor by means of a rope which passes over a hook 6 feet above the floor. If the rope is pulled over the hook at the rate of $\frac{4}{5}$ feet per second, find a general expression for the rate of change of the angle $\theta$ between the rope and the floor (a) as a function of the length $x$ of the rope between the hook and the weight; (b) as a function of the angle $\theta$.

**17.** A police officer in a patrol car is approaching an intersection at 80 feet per second. When he is 210 feet from the intersection a car crosses it, traveling at right angles to the police car path at the rate of 60 feet per second. If the officer focuses his spotlight on this second car, how fast is the light beam turning 2 seconds later, assuming that both vehicles continue at their original rates?

**18.** A ladder 12 feet long leans against a fence 8 feet high, with the lower end on level ground and the upper end projecting over the fence. If the lower end slides away from the fence at the rate of 2 feet per second, find: (a) the rate at which the ladder is rotating when the upper end reaches the top of the fence; (b) the rate at which the ladder begins to rotate as the motion proceeds, the upper end of the ladder now starting to slide down the fence, and the lower end continuing to slide as before.

## 4-6 Simple Harmonic Motion

The name *simple harmonic motion* is used to describe a particular kind of oscillatory motion of a point which at regular intervals moves from one end to the other and then back again on a given line segment. If the line segment, of length $2b$, extends from $x = -b$ to $x = b$ on the $x$-axis, we say that the point $P$ is executing simple harmonic motion on this segment if it is the projection on the $x$-axis of a point $Q$ which is moving around the circle $x^2 + y^2 = b^2$ with constant angular velocity. The relation of $P$ to $Q$ is shown in Fig. 4-26. The angular velocity (in radians per second) of $Q$ is by definition the derivative $d\theta/dt$, where $\theta$ is the angle (measured positively counterclockwise) from the positive $x$-axis to the ray $OQ$.

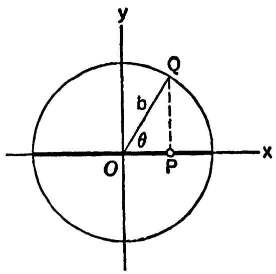



**Fig. 4-26**

If $x$ is the abscissa of $P$, we see that $x = b\cos\theta$. If we denote the constant angular velocity by $\omega$, then

$$\frac{d\theta}{dt} = \omega, \qquad \theta = \omega t + \theta_0,$$

where $\theta_0$ is the value of $\theta$ when $t = 0$. Thus, an expression for $x$ in terms of $t$ is

$$x = b\cos(\omega t + \theta_0). \tag{1}$$

The number $b$ is called the amplitude of the simple harmonic motion. It is the distance from the mid-point to one end of the interval on which $P$ oscillates. The mid-point of the interval is called the *mean position* ("mean" in the sense of "average"). The time $T$ for a complete oscillation is called the *period* of the motion. It is

$$T = \frac{2\pi}{\omega}. \tag{2}$$

The reciprocal of the period is the number of complete oscillations per unit time, and is called the *frequency*.

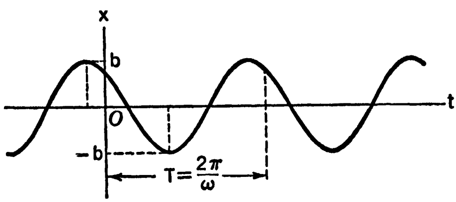



**Fig. 4-27**

The graph of $x$ as a function of $t$ is shown in Fig. 4-27. It is just like the graph of $x = b\cos\omega t$ or $x = b\sin\omega t$ except for a shift along the $t$-axis.

The velocity of $P$ is

$$v = \frac{dx}{dt} = -\omega b\sin(\omega t + \theta_0). \tag{3}$$

From (1) and (3) we readily see that

$$v^2 = \omega^2(b^2 - x^2). \tag{4}$$

This equation shows that the speed of $P$ is greatest (equal to $\omega b$) as $P$

passes through the mean position, and that the speed is 0 at the ends of the interval, where $P$ reverses the direction of its motion.

Many problems about simple harmonic motion may be solved with the aid of equations (2) and (4).

**Example 1:** In a certain simple harmonic motion the moving point has speed 13 feet per second at 3 feet from the mean position, and 5 feet per second at 5 feet from the mean position. Find the amplitude and the period.

*Solution.* In (4) we put $v = 13$, $x = 3$, and then again $v = 5$, $x = 5$. This gives us two equations:

$$169 = \omega^2(b^2 - 9), \qquad 25 = \omega^2(b^2 - 25).$$

The unknowns are $\omega$ and $b$. We subtract the second equation from the first, getting

$$144 = 16\omega^2, \quad \text{or} \quad \omega = 3.$$

Then, substituting back with this value of $\omega$,

$$25 = 9(b^2 - 25), \quad \text{or} \quad b = \frac{5}{3}\sqrt{10}.$$

Thus the amplitude is $\frac{5}{3}\sqrt{10}$ feet and the period is $2\pi/3$ seconds.

For some problems it is necessary to use a formula for $x$ in terms of $t$, either (1) or something equivalent. The formula for $x$ in terms of $t$ is simplest if the instant $t = 0$ occurs when the moving point is either in the mean position or at one end of the interval. If the point is moving to the right through the mean position at $t = 0$, the formula is

$$x = b \sin \omega t. \tag{5}$$

If the point is at $x = b$ when $t = 0$, the formula is

$$x = b \cos \omega t. \tag{6}$$

**Example 2:** In a certain simple harmonic motion of amplitude 16 feet it takes the moving point 6 seconds to travel from the mean position to a distance $8\sqrt{3}$ feet from that position. Find the period.

*Solution.* We can use equation (5), with $b = 16$. Then $t = 6$ gives $8\sqrt{3} = 16 \sin 6\omega$, or $\sin 6\omega = \sqrt{3}/2$. The smallest positive value of $\omega$ satisfying this equation is given by $6\omega = \pi/3$, or $\omega = \pi/18$. Hence the period is $T = 36$ seconds.

If we expand equation (1) we obtain

$$x = b \cos \omega t \cos \theta_0 - b \sin \omega t \sin \theta_0. \tag{7}$$

This has the general form

$$x = A \cos \omega t + B \sin \omega t, \tag{8}$$

where $A$ and $B$ are constants. Conversely, any equation of the form (8), with $A$ and $B$ not both zero, defines a simple harmonic motion. To go from (7) to (8) we have $A = b \cos \theta_0$, $B = -b \sin \theta_0$. To go from (8) to (7) we have

$$b = \sqrt{A^2 + B^2}, \tag{9}$$

$$\sin \theta_0 = \frac{-B}{\sqrt{A^2 + B^2}}, \qquad \cos \theta_0 = \frac{A}{\sqrt{A^2 + B^2}}. \tag{10}$$

In the simple harmonic motion defined by (1) we see from (3) that the acceleration is

$$\frac{d^2x}{dt^2} = -\omega^2 b \cos (\omega t + \theta_0).$$

On comparing with (1) we see that

$$\frac{d^2x}{dt^2} = -\omega^2 x. \tag{11}$$

This formula shows that *the acceleration is directly proportional to $x$ and is opposite in sign to $x$.*

It can be shown that if a point moves on the $x$-axis in such a way that

$$\frac{d^2x}{dt^2} = -kx, \tag{12}$$

where $k$ is a positive constant, then the point moves in simple harmonic motion with $x = 0$ as the mean position. The period will be $2\pi/\sqrt{k}$. The amplitude of the motion is not determined by equation (12), but depends upon the position and velocity of the point at the instant $t = 0$. The justification of the assertions which have just been made can be made a bit later on in this book, after we are in a position to study methods of finding all the functions $x = f(t)$ for which (12) is true. That is, we need to learn how "to solve" the equation (12), which is a particular kind of differential equation. We shall return to this subject later, and learn something about how simple harmonic motion arises in mechanical problems. See § 5-6, Example 4.

## EXERCISES

1. In a certain simple harmonic motion the moving point has speeds of 16 and 20 feet per second at distances of 13 and 5 feet, respectively, from the mean position. (a) Find the amplitude and the period. (b) How long does it take the point to travel 15 feet out from its mean position? (c) How long does it take the speed to fall from its maximum value down to 5 feet per second?

2. In a certain simple harmonic motion the moving point has speeds $v_1$, $v_2$ at the respective distances $x_1$, $x_2$ from the mean position. (a) Assuming

$v_1 > v_2 > 0$, what do you conclude about the relative sizes of $x_1$, $x_2$? (b) Express the period $T$ in terms of $v_1$, $v_2$, $x_1$, $x_2$. (c) Express the amplitude $b$ in terms of $v_1$, $v_2$, $x_1$, $x_2$.

**3.** A point $P$ is moving in simple harmonic motion with period 2 hours. When $P$ is halfway from the mean position to the end of the interval of oscillation, its speed is $30\sqrt{3}$ miles per hour. (a) Find the amplitude of the motion. (b) Find the maximum speed of the point. (c) Find the speed and the distance from the mean position 10 minutes after the point leaves a position of zero velocity. (d) Find the equation expressing the coordinate $x$ of $P$ as a function of $t$ ($x$ in miles, $t$ in hours) if $x = 0$ is the mean position and if $t = 0$ is taken at an instant when $x = b/2$ and $dx/dt > 0$ ($b$ the amplitude). (e) Using the results found in (d) find $x$ and $dx/dt$ when $t = \frac{5}{6}$.

**4.** Consider a simple harmonic motion. (a) In what fraction of the total period does the speed fall from its maximum value to half of this value? (b) When this occurs what fraction of the amplitude is the distance from the mean position at that time?

**5.** (a) In a simple harmonic motion of period $T$ and amplitude $b$, how long does it take for the moving point to move from the mean position $x = 0$ to the position $x = b/2$? (b) What time is required to move from $x = b/2$ to $x = b$?

**6.** A point $Q$ is going at a constant rate around a circle of radius 5 meters. The projection $P$ of $Q$ on a fixed diameter of the circle is traveling $16\pi$ meters per second when it is 3 meters from the center of the circle. (a) How many times per second does $Q$ go around the circle? (b) What is the greatest acceleration of $P$? (c) If $P$ moves on the $x$-axis and if, at $t = 0$, $P$ is at $x = -4$ (meters) going in the negative direction, find the coordinate of $P$ as a function of $t$.

**7.** A point is undergoing simple harmonic motion with frequency 3 oscillations per second. The maximum speed attained by the point is $78\pi$ feet per second. (a) Find the amplitude. (b) How far from the mean position is the point when its speed is $72\pi$ feet per second? (c) If the motion is on the $x$-axis, with mean position at $x = 0$, and if at $t = 0$ we have $x = 5$, $dx/dt < 0$, find $x$ as a function of $t$.

**8.** A point moves on the $x$-axis in simple harmonic motion of period $\pi$. Suppose that $x = 1$ and $dx/dt = 2$ when $t = 0$. (a) Express $x$ as a function of $t$ in the form (8); (b) in the form (1). (c) Find the smallest possible value of $t$ for which $x = 0$. (d) Find the smallest positive value of $t$ for which the velocity is zero.

**9.** A point moves on the $x$-axis in simple harmonic motion of period $4\pi$. (a) If $x = 2$ and $dx/dt = -3$ when $t = 0$, find $x$ as a function of $t$ in the form (8). (b) What is the amplitude of the motion? (c) Find (as a decimal) the smallest positive value of $t$ such that $x = 0$.

**10.** In a certain simple harmonic motion the acceleration is $x'' = -25x$. (a) If $x = -2$ when $t = \pi/2$ and $x = -2\sqrt{3}$ when $t = \pi$, express $x$ as a function of $t$ in the form (8); (b) in the form (1). (c) Find the amplitude and period of the motion. (d) Find $x$ and $x'$ when $t = 0$. (e) Find the smallest positive value of $t$ for which the velocity is zero.

## Review Questions and Problems for Chapters III and IV

### CONCEPTS AND DEFINITIONS

**1.** If $y = f(x)$, express $\Delta y$ in terms of values of $f$, using functional notation.

**2.** Write the definition of the derivative in two different forms, using $y = f(x)$ and the $\Delta$-notation.

**3.** Explain and illustrate the concept of a composite function. If $f$ and $\phi$ are functions, what conditions are needed in order that $f[\phi(x)]$ shall be well defined? Do $f[\phi(x)]$ and $\phi[f(x)]$ define the same function, in general?

**4.** Define what is meant by saying that the graph of $y = f(x)$ ($f$ continuous) is concave upward for $x$ on a certain interval.

**5.** What is a point of inflection?

**6.** Explain the precise meaning of $x^{p/q}$, where $p$ and $q$ are integers, with $q > 0$. Is $x^{p/q}$ a single-valued or a multiple-valued function? Are there limitations on the admissible values of $x$?

**7.** Give the definitions of an ellipse and a hyperbola with reference to two given points as foci.

**8.** If a function $f$ is defined when $a \leq x \leq b$, what is meant by the absolute maximum value of $f$? (Use inequalities to express your answer.) Does such an absolute maximum always exist? Illustrate.

**9.** What is meant by a two-sided relative maximum for a function? Illustrate how such a thing may be the same as the absolute maximum for a given interval, and also how it may differ from the absolute maximum.

**10.** Define the two common systems for measuring angles, and work out the formula for converting from one system to the other.

**11.** Define the first four inverse trigonometric functions and indicate in each case the range of values of the function.

**12.** Define simple harmonic motion:
(a) by relating it to the motion of a point which moves in a circular path;
(b) by indicating how the coordinate of the moving point on the line depends on time;
(c) by a statement about the acceleration.

**13.** Define amplitude, period, mean position, and frequency for a simple harmonic motion.

THEORY

**1.** State fully and prove the theorems about derivatives of sums, products, and quotients of functions. What theorems about limits are used in these proofs?

**2.** Assuming as known the formula $f'(x) = nx^{n-1}$, where $f(x) = x^n$ and $n$ is a positive integer, prove the validity of this same formula when $x \neq 0$ and $n$ is a negative integer. Use the rule for differentiating a quotient.

**3.** State and prove the chain rule for differentiation of composite functions.

**4.** Explain the relationship between the behavior of the derivative $f'(x)$ and the property that the curve $y = f(x)$ is concave upward for $x$ on a certain interval. Hence prove that if $f''(x) > 0$ on the interval, the curve is concave upward.

**5.** Suppose that $f$ has a second derivative at each point of an interval. Does $f''(x_0) = 0$ imply that $(x_0, f(x_0))$ is a point of inflection of the graph of $y = f(x)$? Why is $f''(x_0) = 0$ when $x_0$ corresponds to a point of inflection?

**6.** Prove that $f'(x) = nx^{n-1}$ if $f(x) = x^n$, assuming that $n$ is rational and that $x > 0$.

**7.** Given an equation of the form

$$Ax^2 + By^2 + Cx + Dy + E = 0,$$

what can you say about the locus

(a) If $AB > 0$? (b) If $AB = 0$? (c) If $AB < 0$?
(d) When will there be no locus at all?
(e) When will the locus be a single point?
(f) When will the locus be two intersecting lines?
(g) When will the locus be two parallel lines?
(h) When will the locus be a single line?

**8.** State and prove a theorem, involving the first but not the second derivative, which guarantees the existence of a two-sided relative minimum for a function at $x = x_0$.

**9.** Suppose that $f$ is defined and continuous when $a \leq x \leq b$, differentiable when $a < x < b$, and that it has at most a finite number (i.e., not an infinite number) of two-sided relative maxima and minima in the interval. Explain fully the steps to be taken in finding the absolute maximum and the absolute minimum of the function on the given interval.

**10.** (a) Prove that the sine and cosine functions of $x$ are continuous at $x = 0$. (b) Prove that $\lim_{x\to 0} \frac{\sin x}{x} = 1$, and that $\lim_{x\to 0} \frac{1 - \cos x}{x} = 0$. (c) Derive the formulas for the derivatives of the sine and cosine functions.

**11.** Derive the formulas for the derivatives of $\tan x$, $\text{ctn}\, x$, $\sec x$, and $\csc x$, assuming as known the formulas for the derivatives of $\sin x$ and $\cos x$.

**12.** Derive the formulas for $f'(x)$ and $g'(x)$ if $f(x) = \tan^{-1} x$ and $g(x) = \text{ctn}^{-1} x$.

13. Deduce the formulas $v^2 = p^2(h^2 - x^2)$, $a = -p^2x$, where $v$ and $a$ are the velocity and acceleration, respectively, of a point moving with simple harmonic motion on the $x$-axis, with mean position $x = 0$, amplitude $h$, and period $2\pi/p$.

## PROBLEMS

1. If $y = (ax^2 + bx + c)^{1/2}$, show that $4y^3y'' = 4ac - b^2$.

2. An isosceles triangle of base $2r$ and altitude $h$ is inscribed in a circle of radius $a$. (a) Express the area $A$ of the triangle in terms of $a$ and $h$, after first expressing $r^2$ in terms of $a$ and $h$. (b) Draw the graph of $A$ as a function of $h$ and find the maximum value of $A$. (c) Does the graph have a point of inflection?

3. The triangle and circle of Problem 2 are revolved around the altitude of the triangle, thus generating a cone inscribed in a sphere. (a) Express the lateral area $S$ of the cone as a function of $a$ and $h$. (b) With $a$ fixed, draw the graph of $S$ as a function of $h$, and find the maximum value of $S$. (c) Does the graph have a point of inflection?

4. A small plant produces $x$ units (where $x \geq 5$) of a certain commodity per day, at a total cost $\$C = 4x^3 - 44x^2 + 150x + 144$. The average cost per unit is $A = C/x$, and the marginal cost is defined as $M = dC/dx$. For this case plot $A$ and $M$ as functions of $x$, and show that they intersect at the minimum point of the average cost curve. What is the minimum average cost?

5. Consider the general situation described in Problem 4, but without assuming an explicit formula for $C$. We merely assume that $C$ is a positive, twice differentiable function of $x$ for a certain interval on the positive $x$-axis. (a) Show that $M = A$ when $A$ attains a two-sided relative minimum value. (b) If $M = A$ and $d^2C/dx^2 > 0$, explain carefully how you know that $A$ is at a two-sided relative minimum.

6. The intensity of illumination at any point varies inversely as the square of the distance between the point and the light source and directly as strength of the light source. Two lights, one $r^3$ times as bright as the other, are $c$ feet apart. At what point on the line between the lights is the intensity of illumination least?

7. Two points $A$ and $B$ are diametrically opposite each other on the shore of a circular pond whose radius is 1 mile. A man wishes to go from $A$ to $B$ by swimming from $A$ to a point $P$ on the shore and then walking along the shore from $P$ to $B$. He can swim 2 miles per hour and walk 4 miles per hour. Find the minimum possible time from $A$ to $B$, and also the maximum possible time, under the stated conditions.

8. An airplane $A$ pursues another airplane $B$, which can fly only two thirds as fast as $A$. Both planes remain at the same level. Plane $B$, which was initially 1 mile west of $A$, flies due north, and $A$ continuously heads straight for $B$. Under these conditions the equation of $A$'s path is

$2y = \frac{3}{5}x^{5/3} - 3x^{1/3} + \frac{12}{5}$, where the $x$-axis lies east and west and the origin is at $B$'s initial position. (a) Express $B$'s distance north of $O$ in terms of the $x$-coordinate of $A$. (b) Express the distance between $A$ and $B$ in terms of the $x$-coordinate of $A$. (c) If $B$ flies 200 miles an hour, find the rate of change of $A$'s $x$-coordinate and of the distance between $A$ and $B$ when $x = \frac{1}{8}$.

**9.** One roadway crosses over another at right angles, but on a level 30 feet higher. A car on the upper roadway, going 60 feet per second, passes directly over a car on the lower roadway, going $60\sqrt{3}$ feet per second. How fast are the two cars separating $\frac{1}{3}$ second later?

**10.** A cylinder of radius $b$ with its axis vertical is partly filled with water. A solid right circular cone, its axis vertical, is lowered vertex downward into the water, the vertex descending at the rate of $c$ units per second. Show that the rate at which the water rises in the cylinder is $r^2c/(b^2 - r^2)$, where $r$ is the radius of the base of the submerged part of the cone.

**11.** In a certain type of moving-coil galvanometer a current $i$ produces a deflection of $\theta$ radians, where $i = K\theta/\cos\theta$, $K$ being the galvanometer constant. Graph $i$ as a function of $\theta$ when $0 \leq \theta < \pi/2$. Find the slope at $\theta = 0$. Show that the graph is concave upward.

**12.** A man walks across the diameter of a circular courtyard of radius $b$ feet. A lamp, at one end of the diameter perpendicular to the one on which the man walks, throws his shadow on the wall of the courtyard. Find the speed of the shadow on the wall when the man is $x$ feet from the center of the courtyard, if he walks 6 feet per second.

**13.** An arc light is 24 feet above one side of a street which is 30 feet wide. A man 6 feet tall walks 5 feet per second along the opposite side of the street. When the man is 40 feet along the street from the point opposite the light find (a) how fast his shadow is lengthening; (b) how fast the tip of his shadow is increasing its distance from the point on the ground directly beneath the light; (c) how fast the tip of his shadow is moving.

**14.** A design is made by placing two equal red rectangles at right angles to each other and with a common center so that they just fit inside a white circle of fixed radius $b$. Let $\theta$ be the angle which the diagonal of a rectangle makes with the longer side. Show that the total red area is greatest when $\tan 2\theta = 2$. What is the maximum area?

**15.** A rectangle is inscribed in a circular sector of central angle $2\phi$, a pair of opposite sides of the rectangle being perpendicular to the bisector of the central angle. Prove that the area of the rectangle is greatest when the side having both ends on the circular arc subtends an angle $\phi$ at the center of the circle.

**16.** The clock on a public building has a minute hand 6 feet long and an hour hand 4 feet long. How fast is the distance between the tips of these hands changing at 10 A.M.?

**17.** A conical water glass is to be made so that when a heavy sphere 2 inches in diameter is placed inside and the glass is filled with water, the sphere will barely be submerged. Find the semivertical angle of the cone if the volume of the glass is the least possible.

**18.** (a) For what values of $x$ is $f(x) = (60x - 46x^2 + 12x^3 - x^4)^{1/2}$ defined? (b) For what values of $x$ is $f$ differentiable? (c) Find the absolute maximum and the absolute minimum of $f$. (d) Are there any two-sided relative extrema which are not absolute extrema?

**19.** A ball is tossed straight up. The sun is setting, and the horizontal rays throw the shadow of the ball onto a nearby hemispherical dome, of radius 18 feet. The ball is thrown so that it rises exactly to the height of the top of the dome. (a) Find the speed of the shadow along the surface of the dome as a function of $t$ when $t \geq 0$, if $t = 0$ is taken as the instant at which the ball reaches its highest point. (b) Evaluate at $t = 0$, and note the surprising character of the result.

**20.** A water glass has the shape of a cone of altitude $h$ and semivertical angle $\phi$. The glass is filled with water, and into it is carefully lowered a spherical ball of such size as to cause the greatest possible overflow. Find the radius of the ball.

*CHAPTER V*

# DIFFERENTIALS AND ANTIDERIVATIVES

## 5-1 The Differential of a Function

The concept of a differential is closely related to the concept of a derivative. Suppose $f$ is a function of one independent variable, which we denote by $x$. Let $y$ be the dependent variable, so that $y = f(x)$. Suppose that $f$ is differentiable at $x = x_0$, the value of the derivative being $f'(x_0)$. Then we *define* the differential of $f$ at $x_0$ to be the linear function consisting of all pairs $(\xi, f'(x_0)\xi)$, where $\xi$ can be any real number. Here the independent variable is $\xi$ (Greek xi). If we denote the dependent variable by $\eta$ (Greek eta), then the formula defining the differential is

$$\eta = f'(x_0)\xi. \tag{1}$$

If we regard $\xi$ and $\eta$ as rectangular coordinates with origin at the point $x = x_0$, $y = y_0 = f(x_0)$, with the $\xi$-axis parallel to the $x$-axis and the $\eta$-axis parallel to the $y$-axis, equation (1) is just the equation of the tangent to the curve $y = f(x)$ at the point $(x_0, y_0)$ (see Fig. 5-1). This is evident from the form of the equation (1) and the fact that the slope of the tangent in question is $f'(x_0)$.

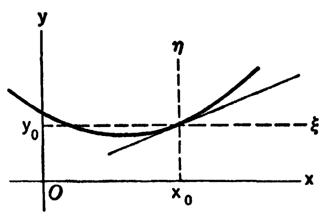


Fig. 5-1

The traditional notation which is used in connection with differentials is due to Leibniz. It is a notation which is convenient in many ways. One of our main purposes in this chapter is to explain this traditional notation and illustrate its uses.

In the customary notation we write $dx$ in place of $\xi$ and $dy$ in place of $\eta$. Thus $dx$ is an independent variable which can be assigned any value; $dy$ is then defined in terms of $x_0$ and $dx$ by the formula $dy = f'(x_0)\,dx$. If we drop the subscript on $x_0$, and regard $x$ as denoting any fixed value of the independent variable associated with the function $f$, then $dy$ is defined in terms of $x$ and $dx$ by the formula

$$dy = f'(x)\,dx. \tag{2}$$

The symbols $dx$ and $dy$ are also called differentials (differential $x$ and differential $y$). If $dx \neq 0$, (2) can be written in the form

$$\frac{dy}{dx} = f'(x), \tag{3}$$

or

$$\frac{\text{differential } y}{\text{differential } x} = \text{derivative of } f \text{ evaluated at } x.$$

Equation (3) looks familiar, of course, for we have been using $dy/dx$ right along as one of the notations for the derivative. The new feature of (3) here is that we have now given meanings to $dx$ and $dy$ as individual things, and the derivative can now be regarded as the quotient of $dy$ by $dx$.

The differential notation is used with letters of any kind to represent the independent and dependent variables. But one must have an understanding in advance as to what function is being considered, and how the letters are being used, before one can apply the definition of the differential.

**Example 1:** The area $A$ of a circle can be considered either a function of the radius $R$ or of the diameter $D$. Find $dA$ for each of these cases.

We know that $A = \pi R^2$, whence

$$\frac{dA}{dR} = 2\pi R, \qquad dA = 2\pi R\,dR. \tag{4}$$

Also, $A = \pi D^2/4$, whence

$$\frac{dA}{dD} = \frac{\pi D}{2}, \qquad dA = \frac{\pi D}{2}\,dD. \tag{5}$$

We are really considering here the differentials of two different functions: the function of $R$ defined by $\pi R^2$, and the function of $D$ defined by $\pi D^2/4$. Consequently the expressions for $dA$ in (4) and (5) are arrived at by entirely separate applications of the definition of the differential. The justification for using the same symbol $dA$ in both cases is this: if we express $D$ as a function of the independent variable $R$ and compute the differential $dD$ accordingly, and if we then substitute these things in (5), the $dA$ of (5) is transformed into the $dA$ of (4). Here are the details:

$$D = 2R, \qquad \frac{dD}{dR} = 2, \qquad dD = 2\,dR,$$

$$\frac{\pi D}{2}\,dD = \pi R \cdot 2\,dR = 2\pi R\,dR.$$

This example illustrates the main reason why the differential notation of Leibniz is so convenient. We state the general principle as a theorem.

THEOREM 5-A. *Suppose that $f$ and $g$ are differentiable functions, with independent variables $x$ and $t$, respectively, and let the values of $g$ lie in the domain of admissible values of $x$. Let $F(t) = f[g(t)]$. If we write $y = f(x)$ and $x = g(t)$, then we also have the formula $y = F(t)$. Now let a value of $t$ be fixed, with $x$ fixed accordingly, and compute $dx$ and $dy$ in terms of $t$ and $dt$ by the general definition of a differential. Then it is still true that $dy = f'(x)\,dx$.*

*Proof.* By definition $dy = F'(t)\,dt$ and $dx = g'(t)\,dt$. But then $f'(x)\,dx = f'[g(t)]g'(t)\,dt$. On the other hand, the chain rule of differentiation (Theorem 3-E) shows that $F'(t) = f'[g(t)]g'(t)$. Hence we see that $f'(x)\,dx = F'(t)\,dt = dy$, as asserted.

In essence, Theorem 5-A says the following: if $y$ is a differentiable function of $x$, then $dy$ divided by $dx$ is always the derivative of $y$ with respect to $x$, no matter what variable is considered as independent when we calculate $dy$ and $dx$.

**Example 2:** If $x = \cos\theta$, $y = \sin\theta$, and $\pi < \theta < 2\pi$, we see that $y < 0$ and $y = -\sqrt{1 - x^2}$. Treating $\theta$ as the independent variable, we have

$$dx = -\sin\theta\,d\theta, \qquad dy = \cos\theta\,d\theta.$$

If we regard $x$ as the independent variable, we have

$$dy = -\frac{-2x}{2\sqrt{1 - x^2}}\,dx = \frac{x\,dx}{\sqrt{1 - x^2}}.$$

In the first case we have

$$\frac{dy}{dx} = \frac{\cos\theta\,d\theta}{-\sin\theta\,d\theta} = \frac{\cos\theta}{-\sin\theta} = \frac{x}{-y} = \frac{x}{\sqrt{1 - x^2}}.$$

This result is consistent with what was obtained when $x$ was regarded as independent.

## 5-2 Standard Differential Formulas

In order to acquire facility with the differential notation, one must practice computing differentials of the various standard types of functions.

For algebraic functions we rely on the rule for dealing with constants and powers, along with the rules for dealing with sums, products, and quotients. The standard formulas are

$$dc = 0, \qquad c \text{ a constant,} \tag{1}$$

$$du^n = nu^{n-1}\,du, \quad n \text{ a constant} \tag{2}$$

$$d(u+v) = du + dv, \tag{3}$$

$$d(uv) = u\,dv + v\,du, \tag{4}$$

$$d\left(\frac{u}{v}\right) = \frac{v\,du - u\,dv}{v^2}. \tag{5}$$

Here $u$ and $v$ represent the dependent variables for any differentiable functions of $x$. Each of these formulas is a direct consequence of a corresponding formula for derivatives. For instance, we know that if $y = u^n$,

$$\frac{dy}{dx} = nu^{n-1}\frac{du}{dx}.$$

When both sides of this equation are multiplied by $dx$, the result is formula (2).

For the trigonometric functions and their inverses we have the formulas:

$$d \sin u = \cos u\,du, \qquad d \cos u = -\sin u\,du, \tag{6}$$

$$d \tan u = \sec^2 u\,du, \qquad d \operatorname{ctn} u = -\csc^2 u\,du, \tag{7}$$

$$d \sec u = \sec u \tan u\,du, \qquad d \csc u = -\csc u \operatorname{ctn} u\,du, \tag{8}$$

$$d \sin^{-1} u = \frac{du}{\sqrt{1-u^2}}, \qquad d \cos^{-1} u = \frac{-du}{\sqrt{1-u^2}}, \tag{9}$$

$$d \tan^{-1} u = \frac{du}{1+u^2}, \qquad d \operatorname{ctn}^{-1} u = \frac{-du}{1+u^2}. \tag{10}$$

**Example 1:** Find $dy$ if $y = (2ax - x^2)^{1/2}$. Using (2), we have

$$dy = \tfrac{1}{2}(2ax - x^2)^{-1/2}\,d(2ax - x^2).$$

Then $d(2ax - x^2) = 2a\,dx - 2x\,dx$, and so

$$dy = \frac{(a - x)\,dx}{(2ax - x^2)^{1/2}}.$$

**Example 2:** Find $dy$ if $y = \sin^2 (x/2)$. Here we use (2) and then the first formula in (6):

$$dy = 2 \sin\frac{x}{2}\,d\left(\sin\frac{x}{2}\right) = 2\sin\frac{x}{2}\cos\frac{x}{2}\,d\left(\frac{x}{2}\right).$$

Since $d\,(x/2) = \frac{1}{2}\,dx$, the final result is

$$dy = \sin\frac{x}{2}\cos\frac{x}{2}\,dx.$$

**Example 3:** Find $dw$ if $w = \tan^{-1}(dy/dx)$, where $y$ is a twice differentiable function of $x$. We use the notation $y'$ for the derivative. Then, using the first formula in (10), we have

$$dw = \frac{1}{1 + y'^2}\,dy'.$$

Now $\qquad \frac{d}{dx}(y') = y''$, and so $dy' = y''\,dx$.

Then $\qquad dw = \frac{y''\,dx}{1 + y'^2}$.

**Example 4:** Use differentials to find an expression for $y'$, supposing that $y$ is a differentiable function of $x$ such that $x^2y^3 - xy + 2 = 0$. Taking the differential of each term in this equation, and using the rules for products, powers, and constants, we obtain

$$x^2 3y^2\,dy + y^3 2x\,dx - x\,dy - y\,dx = 0.$$

Now collect together on the left those terms which contain $dy$, and put the terms in $dx$ on the right:

$$(3x^2y^2 - x)\,dy = (y - 2xy^3)\,dx.$$

Thus finally, $\qquad y' = \frac{dy}{dx} = \frac{y(1 - 2xy^2)}{x(3xy^2 - 1)}$.

As a check against mistakes or omissions when working with differentials, observe that if one term in an equation contains a differential as a factor, then every term must contain a differential.

## EXERCISES

1. Find $dy$ in each case.
   (a) $y = \frac{x}{\sqrt{a^2 + x^2}}$.
   (b) $y = \operatorname{ctn} 2x$.
   (c) $y = \sin^{-1} x^2$.
   (d) $y = \tan^{-1}\sqrt{2x - 1}$.
   (e) $y = x\cos^{-1} x - \sqrt{1 - x^2}$.
   (f) $y = 2x - \sin 2x \cos 2x$.
   (g) $y = \sec^4 3x - \tan^4 3x$.
   (h) $y = \csc^4 5x - \operatorname{ctn}^4 5x$.
2. Find $du$ and simplify the result as much as possible.
   (a) $u = \frac{\sqrt{25 - x^2}}{x} + \sin^{-1}\frac{x}{5}$.
   (b) $u = -\frac{1}{2}x\sqrt{16 - x^2} + 8\sin^{-1}\frac{x}{4}$.
   (c) $u = \frac{\sqrt{x^2 - 9}}{18x^2} + \frac{1}{54}\cos^{-1}\frac{3}{x} \quad (x > 0)$.
   (d) $u = \frac{x - a}{2}\sqrt{2ax - x^2} + \frac{a^2}{2}\sin^{-1}\frac{x - a}{a}$.
3. Find $y'$ by the method of Example 4.
   (a) $16x^3 = 9y^2$.
   (b) $x^2 + y^2 = xy$.
   (c) $x^3 + 3xy^2 + y^3 = 1$.
   (d) $4x^2y - 8xy^2 + 5y^3 = 1$.
   (e) $\cos x + \sin y = 1$.
   (f) $(x^2 + y^2)\tan^{-1}\frac{y}{x} = \pi$.
   (g) $x = 27\csc y + 64\operatorname{ctn} y$.
   (h) $y^2 + x^2y^4 = x^2\sin^2 x$.

4. Express $dy$ in terms of $\cos x$ and $dx$ if $y = \tan^{-1}\left(\frac{5}{3}\tan\frac{x}{2}\right)$. Use half-angle formulas.

5. Express $dy$ in terms of $\sin x$ and $dx$ if $y = \frac{1}{2}\tan^{-1}\left(\frac{5}{4}\tan\frac{x}{2} + \frac{3}{4}\right)$. Use half-angle formulas.

6. Express $dy$ in terms of $\sin x$, $\cos x$, and $dx$ if $y = \frac{1}{ab}\tan^{-1}\left(\frac{b\tan x}{a}\right)$.

## 5-3 Notations for Antiderivatives

The concept of an antiderivative was introduced in § 2-2; the reader should at this time re-examine § 2-2.

There are many interesting problems whose solution involves the finding of an antiderivative of a function which occurs in the problem. We learned something about problems of this sort in § 2-2, but at that stage of our progress the only functions we knew how to deal with were polynomials. Now we know how to differentiate a greater variety of functions, and we are therefore able to find antiderivatives of a more extensive class of functions. In this section we shall make a start at organizing what we know about antiderivatives.

In dealing with antiderivatives of a given function $f$, we shall always suppose that $f$ is defined on some interval of the $x$-axis (perhaps even on the whole $x$-axis). An antiderivative of $f$ is a function $g$, *defined on the same interval as* $f$, and such that $g'(x) = f(x)$ for each $x$ on the interval. If $g$ is one antiderivative of $f$, every possible antiderivative of $f$ is obtained by assigning all possible values to the constant $C$ in the function defined by the expression $g(x) + C$. On this account $g(x) + C$ is often called *the general antiderivative* of $f$. In systematic work with antiderivatives the customary notation for the general antiderivative of $f$ is

$$\int f(x)\,dx.$$

This particular notation has been used since the early days of calculus. It is not logically essential to use the $dx$ in this connection, but we shall see in § 5-4 that the use of the differential notation is convenient in some of the procedures by which we actually find antiderivatives.

Thus far we know the following formulas for antiderivatives. We use the letter $u$ instead of $x$. The particular choice of letter has no essential significance; the choice of $u$ instead of $x$ here is made to fit in with certain later references to these formulas. In each case $C$ represents an arbitrary constant.

I. $\int k\,du = ku + C, \quad k$ a constant.

II. $\int u^n\,du = \dfrac{u^{n+1}}{n+1} + C, \quad n \neq -1.$

III. $\int \cos u\,du = \sin u + C.$

IV. $\int \sin u\,du = -\cos u + C.$

V. $\int \sec^2 u\,du = \tan u + C.$

VI. $\int \csc^2 u\,du = -\operatorname{ctn} u + C.$

VII. $\int \sec u \tan u\,du = \sec u + C.$

VIII. $\int \csc u \operatorname{ctn} u\,du = -\csc u + C.$

IX. $\int \dfrac{du}{\sqrt{1-u^2}} = \sin^{-1} u + C.$

X. $\int \dfrac{du}{1+u^2} = \tan^{-1} u + C.$

In formula IX the variable $u$ is restricted so that $-1 < u < 1$. To check any formula we observe that

$$\int f(u)\,du = g(u) + C \quad \text{is equivalent to} \quad dg(u) = f(u)\,du.$$

There is also a general rule which corresponds to the rule about derivatives expressed in Theorem 3-A: *To find an antiderivative of the sum of two functions, find an antiderivative of each function and add these two antiderivatives.* This rule is sometimes expressed in the form

$$\int [f_1(x) + f_2(x)]\,dx = \int f_1(x)\,dx + \int f_2(x)\,dx, \tag{1}$$

but the interpretation of this formula calls for some special comment, in view of the earlier statement that $\int f(x)\,dx$ denotes the general antiderivative of $f$. A "general antiderivative" is not a single function, but a family of functions obtained by adding all possible constants to some particular antiderivative. How, therefore, are we to interpret (1), which in some way is supposed to express, not an equality of two functions, but an equality of two families of functions? There are various ways of giving a formal interpretation of (1); some ways involve more mathematical sophistication than others. We shall put the matter as follows. Formula (1) is a short way of summing up two statements:

(a) *Every* antiderivative of $f_1 + f_2$ can be expressed by adding *some* antiderivative of $f_1$ and *some* antiderivative of $f_2$;

(b) the sum of *any* antiderivative of $f_1$ and *any* antiderivative of $f_2$ is *some* antiderivative of $f_1 + f_2$.

Other formulas which occur later and involve the symbol $\int$ more than once, or which involve the symbol $\int$ and the $C$ denoting an arbitrary constant, are to be interpreted in a similar manner as statements asserting the equality of two families of functions.

Formula (1) can be extended to more than two terms on the right, and the extension indicates an actual procedure for finding antiderivatives where sums and differences are involved.

**Example 1:**

$$\int (15x^4 + 2\cos x + 3\sec^2 x)\,dx = \int 15x^4\,dx + \int 2\cos x\,dx + \int 3\sec^2 x\,dx$$
$$= 3x^5 + 2\sin x + 3\tan x + C.$$

Observe that only one arbitrary constant is needed to express the general antiderivative, even though the symbol $\int$ was used several times.

For constant factors $k$ we have the formula

$$\int k\,f(x)\,dx = k\int f(x)\,dx, \tag{2}$$

which expresses the rule that *to find an antiderivative of $kf(x)$, find an antiderivative of $f(x)$ and multiply it by $k$.*

**Example 2:**

$$\int 2\sin x\,dx = 2\int \sin x\,dx = 2(-\cos x) + C.$$

## 5-4 Antiderivatives by Substitution

To find an antiderivative of a given function we must be able by some means to recognize $f(x)$ as the derivative of another function; or what is equivalent, we must recognize $f(x)\,dx$ as the differential of this other function. In many cases such recognition may be made by using a substitution to reduce the expression $f(x)\,dx$ to a form which is more readily recognizable as the differential of a known function.

**Example 1:** Find $\int \frac{x\,dx}{\sqrt{4+x^2}}$. If we let $u = 4 + x^2$, then $du = 2x\,dx$, and so

$$\frac{x\,dx}{\sqrt{4+x^2}} = \frac{1}{2}\frac{du}{\sqrt{u}} = \frac{1}{2}u^{-1/2}\,du = \frac{1}{2}d\left(\frac{u^{1/2}}{\frac{1}{2}}\right) = d(\sqrt{4+x^2}).$$

Hence
$$\int \frac{x\,dx}{\sqrt{4+x^2}} = \sqrt{4+x^2} + C.$$

This example illustrates the method of substitution. The general principle may be stated as follows.

THEOREM 5-B. *Let it be required to find $\int f(x)\,dx$, and suppose that $h$ is a differentiable function of $x$, defined on the same interval as $f$, such that when we substitute $u = h(x)$, the expression $f(x)\,dx$ becomes $\phi(u)\,du$, where $\phi$ is a function of $u$ for which we can find an antiderivative. If we then write down the explicit expression for $\int \phi(u)\,du$ and in it replace $u$ by $h(x)$, we shall have $\int f(x)\,dx$.*

*Proof.* Let $\Phi'(u) = \phi(u)$, so that $\Phi$ is an antiderivative of $\phi$, and

$$\int \phi(u)\,du = \Phi(u) + C.$$

We have to show that

$$\int f(x)\,dx = \Phi[h(x)] + C, \tag{1}$$

or, equivalently, that

$$f(x)\,dx = d\Phi[h(x)]. \tag{2}$$

Now, if $u = h(x)$,

$$d\Phi(u) = \Phi'(u)\,du = \phi(u)\,du = f(x)\,dx.$$

Hence we see that (2) is true. This proves (1) and establishes the truth of the theorem.

The ability to detect a good substitution for a particular problem must be developed by practice. One must gradually become familiar with various types of problems, learning by observation and example the kinds of substitutions that are appropriate for each type.

Linear substitutions, such as $u = 2x$, or $v = -3x + 5$, are often useful.

**Example 2:** Find $\int (3 - 7x)^{3/2}\,dx$. Let $u = 3 - 7x$, $du = -7\,dx$. Then $dx = -du/7$, and

$$\int (3 - 7x)^{3/2}\,dx = -\frac{1}{7}\int u^{3/2}\,du = -\frac{1}{7}\cdot\frac{u^{5/2}}{\frac{5}{2}} + C = -\frac{2}{35}(3 - 7x)^{5/2} + C.$$

**Example 3:** Find $\int \cos 3x\,dx$. Let $u = 3x$. Then

$$\int \cos 3x\,dx = \frac{1}{3}\int \cos u\,du = \frac{1}{3}\sin u + C = \frac{1}{3}\sin 3x + C.$$

After a little practice, one gets accustomed to making this kind of substitution mentally, and the steps need not all be written down. The solution of Example 3 may be written in the form

$$\int \cos 3x\,dx = \frac{1}{3}\int \cos 3x\,d(3x) = \frac{1}{3}\sin 3x + C.$$

To find antiderivatives of the types

$$\int (x^2 \pm a^2)^n x\,dx, \qquad \int (a^2 - x^2)^n x\,dx,$$

where $n$ need not be an integer, one may substitute, respectively,

$u = x^2 \pm a^2$ and $u = a^2 - x^2$. The presence of the combination $x\,dx$ is important, for if we had merely $dx$, instead of $x\,dx$, the expression of $dx$ in terms of $u$ and $du$ might lead to complications.

**Example 4:** In $\int x\sqrt{16 - x^2}\,dx$ let $u = 16 - x^2$. Then $du = -2x\,dx$, so that $x\,dx = -du/2$, and

$$\int x\sqrt{16 - x^2}\,dx = -\frac{1}{2}\int u^{1/2}\,du = -\frac{1}{2}\frac{u^{3/2}}{\frac{3}{2}} + C$$
$$= -\frac{1}{3}(16 - x^2)^{3/2} + C.$$

For an alternative substitution in problems of this type one may use $u^2$ in place of $u$; i.e., we can let $u = \sqrt{x^2 \pm a^2}$ or $u = \sqrt{a^2 - x^2}$. This procedure is preferred by some people because it frequently avoids the use of fractional exponents.

We can deal with expressions of the type

$$\int \sin^m ax \cos ax\,dx, \qquad \int \cos^n ax \sin ax\,dx$$

by letting $u = \sin ax$ in the first case and $u = \cos ax$ in the second. The essential thing here is to have the combination which gives us a power of $u$ times $du$, with perhaps a constant factor.

**Example 5:** In $\int \cos^3 6x \sin 6x\,dx$ let $u = \cos 6x$. Then $du = -\sin 6x\,d(6x)$ $= -6\sin 6x\,dx$, $\sin 6x\,dx = -du/6$, and

$$\int \cos^3 6x \sin 6x\,dx = -\frac{1}{6}\int u^3\,du = -\frac{1}{24}u^4 + C$$
$$= -\frac{1}{24}\cos^4 6x + C.$$

## EXERCISES

**1.** Find antiderivatives in each case by an appropriate substitution.

(a) $\int \sqrt{1 - 2x}\,dx$.

(b) $\int \sin 5x\,dx$.

(c) $\int \frac{dx}{\sqrt{2 - 3x}}$.

(d) $\int \frac{x\,dx}{(9 - x^2)^{3/2}}$.

(e) $\int \cos 2x \sin^2 2x\,dx$.

(f) $\int \cos\frac{1 - x}{2}\,dx$.

(g) $\int \sqrt{\cos\frac{x}{2}}\sin\frac{x}{2}\,dx$.

(h) $\int \frac{dx}{\cos^2 4x}$.

(i) $\int \frac{\cos 3x}{\sin^2 3x}\,dx$.

(j) $\int \frac{du}{\sin^2(2 - u)}$.

(k) $\int \frac{dy}{\sqrt{1 - (1 - 3y)^2}}$.

(l) $\int \frac{dt}{1 + (3t - 4)^2}$.

**2.** Find each antiderivative, using an appropriate substitution. Check by differentiation.

(a) $\int \frac{dx}{(5x-2)^{5/2}}$.

(b) $\int \frac{x\,dx}{(x^2+25)^5}$.

(c) $\int \frac{\sin\theta}{\cos^3\theta}\,d\theta$.

(d) $\int \frac{dx}{1+9x^2}$.

(e) $\int \frac{dx}{\csc 7x}$.

(f) $\int x(b^2-x^2)^{-4}\,dx$.

(g) $\int \frac{x\,dx}{\sqrt{(x^2-16)^3}}$.

(h) $\int \frac{dx}{\sqrt{1-25x^2}}$.

(i) $\int \csc^2 3y\,dy$.

(j) $\int x(a^2-x^2)^{5/2}\,dx$.

(k) $\int \frac{dx}{\sqrt[3]{(4-x)^2}}$.

(l) $\int \tan\frac{u}{2}\sec\frac{u}{2}\,du$.

**3.** Find each of the following antiderivatives in two ways: once by the method illustrated in Example 4, and once by the alternative substitution suggested right after Example 4, with $u^2$ in place of $u$.

(a) $\int x\sqrt{x^2-a^2}\,dx$.

(b) $\int \frac{x}{\sqrt{a^2-x^2}}\,dx$.

(c) $\int \frac{x}{(a^2+x^2)^{3/2}}\,dx$.

(d) $\int \frac{x\,dx}{(a^2-x^2)^2}$.

(e) $\int x(x^2-a^2)^{3/2}\,dx$.

(f) $\int \frac{x\,dx}{(x^2+a^2)^3}$.

**4.** Find each antiderivative. Check by differentiation.

(a) $\int \left(\frac{4}{\cos^2 x}+\frac{7}{x^2}\right)dx$.

(b) $\int \left(1+\frac{x}{4}\right)^{5/2}dx$.

(c) $\int \frac{3-\sin 2x}{\cos^2 2x}\,dx$.

(d) $\int \frac{1-2\cos 3x}{\sin^2 3x}\,dx$.

(e) $\int \frac{1-x^2}{1-x^4}\,dx$.

(f) $\int \frac{2-3x}{\sqrt{x}}\,dx$.

(g) $\int \frac{x\,dx}{(2x^2+1)^2}$.

(h) $\int \frac{x\,dx}{(49-25x^2)^{1/2}}$.

**5.** Find each antiderivative. In some of these problems the appropriate substitution is different from any of those which have been illustrated.

(a) $\int \frac{\sin x}{(1-\cos x)^2}\,dx$.

(b) $\int \frac{\cos x}{(2+3\sin x)^2}\,dx$.

(c) $\int \frac{\cos x}{1+\sin^2 x}\,dx$.

(d) $\int (1+x^3)^{1/2}x^2\,dx$.

(e) $\int \frac{x^3}{\sqrt{1+x^4}}\,dx$.

(f) $\int \tan^2 x\sec^2 x\,dx$.

(g) $\int x\sin x^2\,dx$.

(h) $\int \frac{dx}{16+9x^2}$.

(i) $\int \frac{x\,dx}{1+x^4}$.

(j) $\int \frac{\sec^4 x}{\csc x}\,dx$.

## 5-5 Some Standard Formulas

For practical purposes it is best to replace the standard formulas IX and X in § 5-3 by slightly more general formulas as follows:

IX′. $$\int \frac{dx}{\sqrt{a^2 - x^2}} = \sin^{-1}\frac{x}{a} + C,$$

X′. $$\int \frac{dx}{a^2 + x^2} = \frac{1}{a}\tan^{-1}\frac{x}{a} + C.$$

Here it is assumed that $a$ is a positive constant. To deduce IX′ from IX, make the substitution $x = au$ in the left side of IX′. The result is

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \int \frac{a\,du}{\sqrt{a^2(1 - u^2)}} = \int \frac{du}{\sqrt{1 - u^2}}$$

$$= \sin^{-1} u + C = \sin^{-1}\frac{x}{a} + C.$$

The derivation of X′ from X is similar. Hereafter we shall use IX′ and X′ as standard formulas.

We often use IX′ or X′ after making a slight preliminary transformation in a problem.

**Example 1:** Find $\int \frac{dx}{\sqrt{5 - 4x^2}}$. We can either write

$$\sqrt{5 - 4x^2} = \sqrt{4\left(\frac{5}{4} - x^2\right)} = 2\sqrt{\left(\frac{\sqrt{5}}{2}\right)^2 - x^2}$$

and use IX′, or we can let $2x = u$ and obtain

$$\int \frac{dx}{\sqrt{5 - 4x^2}} = \frac{1}{2}\int \frac{du}{\sqrt{5 - u^2}} = \frac{1}{2}\sin^{-1}\frac{u}{\sqrt{5}} + C = \frac{1}{2}\sin^{-1}\frac{2x}{\sqrt{5}} + C.$$

The result is the same either way.

The following formula is often used in connection with problems about circles and ellipses:

XI. $$\int \sqrt{a^2 - x^2}\,dx = \frac{a^2}{2}\sin^{-1}\frac{x}{a} + \frac{x}{2}\sqrt{a^2 - x^2} + C.$$

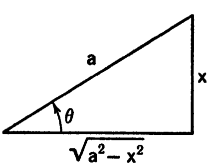


**Fig. 5-2**

The derivation of this formula may be made as follows. We let $\theta = \sin^{-1}(x/a)$, so that $x = a\sin\theta$, $dx = a\cos\theta\,d\theta$. This substitution is suggested by thinking of the right triangle labeled as in Fig. 5-2. Since $\sqrt{a^2 - x^2} = a\cos\theta$, we see that

$$\int \sqrt{a^2 - x^2}\,dx = a^2\int \cos^2\theta\,d\theta.$$

Now the half-angle formulas

$$2\cos^2\theta = 1 + \cos 2\theta, \qquad \sin 2\theta = 2\sin\theta\cos\theta$$

prove to be convenient:

$$\int \sqrt{a^2 - x^2}\,dx = \frac{a^2}{2}\int (1 + \cos 2\theta)\,d\theta = \frac{a^2}{2}\int d\theta + \frac{a^2}{2}\int \cos 2\theta\,d\theta$$

$$= \frac{a^2\theta}{2} + \frac{a^2}{4}\sin 2\theta + C$$

$$= \frac{a^2\theta}{2} + \frac{a^2}{2}\sin\theta\cos\theta + C.$$

On expressing this result in terms of $x$, we obtain XI, because $a^2 \sin\theta\cos\theta = x\sqrt{a^2 - x^2}$.

We are now able to calculate the area enclosed by an ellipse.

**Example 2:** Find the first quadrant portion of the area enclosed by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

We use the method of § 2-7. If $S(x)$ is the area shown with shading in Fig. 5-3, we know that

$$\frac{dS}{dx} = y = \frac{b}{a}\sqrt{a^2 - x^2}.$$

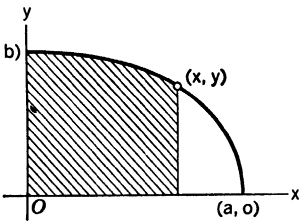



**Fig. 5-3**

Now XI enables us to express $S$ as an antiderivative:

$$S(x) = \frac{ab}{2}\sin^{-1}\frac{x}{a} + \frac{bx}{2a}\sqrt{a^2 - x^2} + C, \tag{1}$$

for some suitable value of $C$. To evaluate $C$, we know that $S = 0$ when $x = 0$; putting these values in (1), we see that $0 = S(0) = C$. Finally, putting $x = a$, we see that the required first quadrant area is

$$S(a) = \frac{ab}{2}\sin^{-1} 1 + 0 = \frac{\pi ab}{4}.$$

The entire area bounded by the ellipse is 4 times as great, or $\pi ab$.

## EXERCISES

**1.** Find each antiderivative, using one or the other of the procedures suggested in Example 1. It is well to practice both procedures.

(a) $\int \frac{dx}{\sqrt{16 - 9x^2}}$. (d) $\int \frac{dx}{8 + 9x^2}$.

(b) $\int \frac{dx}{25 + 4x^2}$. (e) $\int \frac{dx}{\sqrt{4 - 3x^2}}$.

(c) $\int \frac{dx}{\sqrt{8 - 25x^2}}$. (f) $\int \frac{dx}{3x^2 + 4}$.

2. (a) Derive a general formula for $\int \frac{dx}{\sqrt{b^2 - a^2x^2}}$, where $b$ and $a$ are positive. Check it by differentiation.

   (b) Proceed as in (a) with $\int \frac{dx}{(b^2 + a^2x^2)}$.

3. Find the area inside the circle $x^2 + y^2 = 8$ and between the lines $x = \sqrt{2}$, $x = 2$.

4. Find the area inside the ellipse $\frac{x^2}{16} + \frac{y^2}{9} = 1$ and between the lines $x = -2$, $x = 2\sqrt{3}$.

5. Find each antiderivative, and check the results by differentiation.

   (a) $\int \sqrt{16 - 9x^2}\, dx$. (b) $\int \sqrt{b^2 - a^2x^2}\, dx \quad (a, b > 0)$.

## 5-6 More About Acceleration

Consider the motion of a particle on a straight line, which we take to be the $x$-axis. We know that the velocity and acceleration of the particle are, respectively,

$$v = \frac{dx}{dt} \quad \text{and} \quad a = \frac{dv}{dt} = \frac{d^2x}{dt^2}. \tag{1}$$

Another useful expression for the acceleration is found by taking advantage of the fact that derivatives can be written as quotients of differentials. We multiply and divide by $dx$, and shift the positions of the $dx$'s to suit our convenience:

$$a = \frac{dv}{dt} = \frac{dv}{dt}\frac{dx}{dx} = \frac{dx}{dt}\frac{dv}{dx} = v\,\frac{dv}{dx}. \tag{2}$$

**Example 1:** A ball, rolling up a certain incline, is slowed down at the rate of 9 feet per second per second. If the ball is moving 12 feet per second when it passes a certain point, how far does it roll before it stops and begins to roll down?

This problem could be solved by the methods used in § 2-3. But we shall solve it by using the formula for acceleration in (2). We take the $x$-axis to extend up the incline, with $x = 0$ at the point where the ball is moving 12 feet per second. Then $a = -9$, so

$$v\,\frac{dv}{dx} = -9, \quad \text{or} \quad v\,dv = -9\,dx.$$

Passing to antiderivatives, we conclude that

$$\frac{v^2}{2} = -9x + C.$$

To find the value of $C$ we put $x = 0$, $v = 12$:

$$\frac{144}{2} = C, \qquad C = 72.$$

Thus the general formula connecting $v$ and $x$ is

$$v^2 = -18x + 144.$$

This formula shows that $x = 8$ when $v = 0$. This means that the ball stops after going 8 feet up the incline.

The expression for the acceleration in (2) is especially useful when the acceleration is known to depend in some specified way on $x$ or $v$. We shall consider several interesting problems of this type. These problems arise naturally through the use of Newton's second law of motion to determine the acceleration.

Newton's law asserts that when a particle of mass $m$ is moved by a force $F$, the product of the mass and its acceleration is a constant multiple of the force. That is,

$$ma = kF. \tag{3}$$

The value of $k$ depends only on the units used for mass, distance, time, and force, and not on the particular problem under consideration. It is convenient to have units such that $k = 1$. Such is the case, for example, in each of the systems indicated in the adjoining table. If the pound instead of the poundal is used for the unit of force, the corresponding value of $k$ is approximately 32 (the same as the acceleration due to gravity in the British system).

| Mass unit | Distance unit | Time unit | Force unit |
|---|---|---|---|
| gram | centimeter | second | dyne |
| kilogram | meter | second | newton* |
| slug* | foot | second | pound |
| pound | foot | second | poundal * |

* 1 newton = $10^5$ dynes; 1 slug = $g$ pounds and 1 poundal = $1/g$ pound, where $g = 32$, approximately.

**Example 2:** Suppose a particle moves on the positive $x$-axis under the influence of a force which attracts the particle toward the origin, the magnitude of the force being inversely proportional to the square of the distance from the origin. Suppose the particle has velocity $v_0$ when $x = x_0$. Find the general formula connecting $v$ and $x$.

From Newton's law we see that

$$v\frac{dv}{dx} = -\frac{k}{x^2}, \quad \text{or} \quad v\,dv = -kx^{-2}\,dx, \tag{4}$$

where $k$ is a positive proportionality constant, not having the same significance as the $k$ in (3). The acceleration is negative, because the attraction is toward the origin. From (4) we conclude that

$$\frac{v^2}{2} = -k\frac{x^{-1}}{-1} + C = \frac{k}{x} + C. \tag{5}$$

We put $v = v_0$, $x = x_0$, and solve for $C$. On putting the value of $C$ back in (5), we find

$$v^2 = v_0^2 - \frac{2k}{x_0} + \frac{2k}{x}. \tag{6}$$

This is the required formula. For a discussion of special cases of this problem and of the significance of the sign of the quantity $v_0^2 - (2k/x_0)$, see some of the exercises.

**Example 3:** Experiments show that when an object moves through water at moderate speed, it is in many cases a satisfactory approximation of the situation to say that the magnitude of the resistance offered by the water is directly proportional to some power of the velocity. Suppose, for a certain boat, that the resistance of the water is proportional to the three-halves power of the velocity, and that the acceleration due to the resistance is $-3$ feet per second per second when the boat is going 36 feet per second. With all power off, how far does the boat go while the speed is dropping from 36 to 16 feet per second? How long does this take?

We know from Newton's law that $a = -kv^{3/2}$, and we are given that $a = -3$ when $v = 36$. Therefore $-3 = -k(6)^3 = -216k$, or $k = \frac{1}{72}$. Then

$$a = v\frac{dv}{dx} = -\frac{1}{72}v^{3/2}, \quad \text{or} \quad 72v^{-1/2}\,dv = -dx.$$

Proceeding to antiderivatives, we obtain

$$72\frac{v^{1/2}}{\frac{1}{2}} = 144v^{1/2} = -x + C_1.$$

We assume that $x = 0$ and $t = 0$ when $v = 36$; this gives $C_1 = 144(6) = 864$. Thus

$$144v^{1/2} = -x + 864.$$

In this result we put $v = 16$ and find $x = 288$. This means that the boat goes 288 feet while the speed is falling from 36 to 16 feet per second. To find the time required for this to occur, we proceed as follows:

$$v = \frac{dx}{dt} = \left(\frac{864 - x}{144}\right)^2, \qquad \frac{(144)^2}{(x - 864)^2}\,dx = dt.$$

Then, going to antiderivatives, we have

$$(144)^2\frac{(x - 864)^{-1}}{-1} = t + C_2. \tag{7}$$

Putting $x = 0$, $t = 0$, we obtain $C_2 = 24$. We put this back in (7), set $x = 288$, and solve for $t$. The result is $t = 12$. This means that the boat takes 12 seconds to travel the 288 feet.

**Example 4:** In a variety of physical problems the motion of a particle on a straight line (the $x$-axis) is such that the acceleration is a constant negative multiple of the coordinate $x$, so that we can write

$$\frac{dv}{dt} = \frac{d^2x}{dt^2} = v\frac{dv}{dx} = -kx, \tag{8}$$

where $k > 0$. This kind of motion can be realized if the motion is produced by a force which is directly proportional to the distance from the origin to the particle, the force always being *toward the origin.* Various mechanical devices with springs or stretched rubber bands can be devised to produce this kind of a force situation. Our concern here is to demonstrate that acceleration of the type specified in (8) always leads to simple harmonic motion. Simple harmonic motion was discussed previously in § 4-6.

Writing (8) in the form $v\,dv = -kx\,dx$, we proceed by antidifferentiation to obtain

$$\frac{v^2}{2} = -\frac{kx^2}{2} + C.$$

If we let $b$ be that positive value of $x$ which makes $v = 0$ in this formula, we find that $C = kb^2/2$, whence

$$v^2 = \left(\frac{dx}{dt}\right)^2 = k(b^2 - x^2).$$

Next, on taking the square root, we have

$$\frac{dx}{\sqrt{b^2 - x^2}} = \pm\sqrt{k}\,dt.$$

The choice of sign here will depend on whether the velocity is positive or negative at the moment. If we assume for definiteness that $v > 0$, then from the standard formula IX′ in § 5-5 we see that

$$\sin^{-1}\frac{x}{b} = \sqrt{k}\,t + C_1, \quad \text{or} \quad x = b\sin(\sqrt{k}\,t + C_1),$$

where $C_1$ is some constant. We recognize from this formula that the particle is moving with simple harmonic motion. The amplitude is $b$ and the period is $2\pi/\sqrt{k}$.

## EXERCISES

1. A ball is rolled across a level field, its initial velocity being 25 feet per second. (a) If friction slows the ball at the rate of 10 feet per second per second, how far will the ball roll? (b) Express the distance rolled (at any time prior to stopping) as a function of the velocity at that time. (c) Express the velocity at any moment as a function of the distance rolled from the initial point. From this, by antidifferentiation, obtain the time to roll a given distance as a function of that distance.

2. The driver of an automobile finds that he can increase his speed from 15 feet per second to 60 feet per second while going a distance of 300 feet. (a) What uniform acceleration is required to accomplish this result? (b) If the car travels $x$ feet in $t$ seconds after the moment when the speed was 15 feet per second, express $dx/dt$ as a function of $x$. Then, by antidifferentiation, obtain $t$ as a function of $x$.

3. A train is going 60 miles per hour when the brakes are slammed on. The train comes to a stop after going $\frac{1}{4}$ mile. (a) Find the deceleration, in feet per second per second, assuming it is constant. (b) How far does the train go while the speed is being reduced to 30 miles per hour?

4. Suppose a point moves on the $x$-axis with constant acceleration $k$, with $x = 0$ and velocity $v = v_0$ when $t = 0$. (a) Assuming $v_0 > 0$ and $k < 0$, find the value of $x$ for which $v = 0$. (b) Find $k$ in terms of $v_0$, $v_1$, and $x_1$ if $v = v_1$ when $x = x_1$. (c) Assuming $v_0 > 0$ and $v > 0$, express $v$ as a function of $x$, and then by antidifferentiation express $t$ as a function of $x$.

5. Suppose the $x$-axis sticks out of the earth, with the origin at the center of the earth. A particle of mass $m$ on the $x$-axis and outside the earth is attracted toward the origin by a force $cmM/x^2$, where $M$ is the mass of the earth and $c$ is a constant depending only on the units used. At the surface of the earth this attraction is $mg$, the weight of the particle. Hence, if the radius of the earth is $R$, we find that $cM = gR^2$. Therefore $cmM/x^2 = mgR^2/x^2$. In the following problems take $R = 4000$ miles. When distances are in miles and times are in seconds the value of $g$ is $\frac{1}{165}$.
(a) If air resistance and the gravitational influences of the moon and other heavenly bodies were negligible, with what speed would a projectile have to be fired straight up from the earth in order to rise 4000 miles before stopping? In order to rise 40,000 miles? In order to keep going forever? Express all of your answers as multiples of $\sqrt{gR}$ before computing them.
(b) If a rocket could propel itself vertically to a height of 200 miles before exhausting its fuel, what velocity should it then have in order to rise 3800 miles more? In order to keep going forever? Express your answers in terms of $g$ and $R$ before doing the final computations.
(c) The mass of the moon is about $\frac{1}{81}$ that of the earth. Show that the point between the earth and the moon, where the two exert equal (but opposite) gravitational pulls on a particle, is $\frac{9}{10}$ of the way from the center of the earth to the center of the moon. Find the acceleration of a particle at distance $x$ from the center of the earth on a line between the earth and moon. Take the moon's distance $D$ from the earth (center to center) to be 237,000 miles. Then, by antidifferentiation, find the velocity of the particle, assuming it is a projectile fired with velocity $v_0$ from the earth's surface, straight toward the moon. The projectile will reach the moon if $v_0$ is large enough to bring the projectile to the point $x = \frac{9}{10}D$ with a positive velocity. Show that the required $v_0$ is nearly 99 per cent of the initial velocity the projectile would have to have to keep going forever if the gravitational influence of the moon were ignored.

**6.** In Example 2 of the text let $p_0 = (2k/x_0) - v_0^2$, and consider the sign of $p_0$. Suppose that $v_0 > 0$, so that the particle moves in the positive direction as it leaves the position $x = x_0$. Show that there are two cases: Case I, in which the particle comes to a stop at a point on the positive $x$-axis and then moves back toward the origin, and Case II, in which the particle moves always in the positive direction. Case I is characterized by $p_0 > 0$, and Case II by $p_0 \leq 0$. In Case II the particle approaches a limiting velocity $\sqrt{-p_0}$ as $x$ increases indefinitely.

**7.** If $dv/dt = -4v^{1/2}$, and if $x = 0$ and $v = 64$ when $t = 0$, find (a) $v$ in terms of $t$, (b) $x$ in terms of $t$. Then use $dv/dt = v\,dv/dx$ to find (c) $v$ in terms of $x$, and (d) $t$ in terms of $x$. (e) What are the values of $t$ and $x$ when $v = 0$? (f) What are the values of $v$ and $t$ when $x = 36$?

**8.** Suppose $dv/dt = -kv^{1/2}$, and that $v = 81$ and $x = 0$ when $t = 0$. Suppose also that $v = 0$ when $t = 6$. Find (a) $v$ in terms of $t$ and $k$; (b) the value of $k$; (c) $x$ in terms of $t$. Then use $dv/dt = v\,dv/dx$ to find (d) $v$ in terms of $x$; (e) $t$ in terms of $x$. (f) What are the values of $t$ and $x$ when $v = 0$? (g) What are the values of $v$ and $t$ when $x = 114$?

**9.** Assume a law of motion $v\,dv/dx = -kv^{3/2}$ for a boat, much as in Example 3 of the text. Suppose that $v = 25$ when $t = 0$ and $x = 0$, and that $x = 100$ when $v = 16$. (a) Find $v$ in terms of $k$ and $x$. (b) Find the value of $k$. (c) Find $x$ when $v = 9$. (d) Find the relation between $x$ and $t$. (e) What do $x$ and $v$ approach, respectively, as $t \to \infty$?

**10.** A delayed-action bomb, of a certain size and shape, is retarded, when it strikes the earth, at a rate proportional to the square root of the velocity. If for an impact speed of 225 feet per second the bomb will penetrate to a depth of $3\frac{3}{8}$ feet, find (a) the time required for the bomb to come to rest, and (b) the corresponding time and depth of penetration for an identical bomb, if the speed at impact is 400 feet per second.

**11.** The acceleration of a particle moving on the $x$-axis is given as $-kx$, where $k$ is a positive constant. It is given that the velocity is 9 when $x = 0$ and 6 when $x = 3$. (a) Find the value of $k$ and the general relation between $v$ and $x$. (b) Find the amplitude and period of the simple harmonic motion. (c) Express $x$ as a function of $t$, assuming that $t = 0$ when $x = 0$ and $v = 9$.

**12.** Imagine a tunnel of small diameter to be bored through the earth from one side to the other, directly through the center. Then, if the earth were of uniform density, the effect of gravitational attraction would be such that a particle in the tunnel would be attracted toward the center of the earth by a force proportional to the distance from the center. The constant of proportionality can be evaluated by using the known magnitude of the force when the particle is at one end of the tunnel. Show that if the particle were dropped into the tunnel at one end, it would traverse the tunnel from one end to the other with simple harmonic motion. Find the period of the motion and the speed of the particle at the center of the earth. Denote the radius of the earth by $R$, taking $R = 4000$ miles. The value of $g$ for miles and seconds as units is $\frac{1}{165}$.

## 5-7 Parametric Representation

There are cases in which a curve is more easily and naturally described, not by giving an equation which the coordinates $(x, y)$ of a point on the curve must satisfy, but by giving *two equations*, one equation expressing $x$ as a function of an auxiliary variable, and another equation expressing $y$ as a function of this same auxiliary variable. The auxiliary variable is usually called a *parameter*. The description of the curve by means of two equations in this way is called a *parametric representation* of the curve.

**Example 1:** Consider the equations

$$x = \frac{1}{10}(t + t^3), \qquad y = \frac{1}{36}(2t + t^5). \tag{1}$$

Here the parameter is $t$. Let us see what we can find out about the way in which the point $(x, y)$ moves as $t$ varies. Some points can be located by calculating $x$ and $y$ for several values of $t$; a few such calculations are worth while to give us something definite to look at. But in general it is more profitable to study how $x$ and $y$ vary as $t$ varies, instead of merely plotting points. In the present case we observe at the outset that we can write

$$x = \frac{t}{10}(1 + t^2), \qquad y = \frac{t}{36}(2 + t^4). \tag{2}$$

From these equations it is clear that $x$ and $y$ have the same sign as $t$; also, if we change the sign but not the magnitude of $t$, then $x$ and $y$ likewise change in sign but not in magnitude. Hence it will be sufficient to investigate the situation when $t \geq 0$. Now, as $t$ increases, it is clear from equations (2) that $x$ and $y$ both increase. If we make a table of values, say for $t = 0, 1, 2, 3$, we get several points and we can use the foregoing information to give us some confidence in drawing the curve which is represented parametrically by equations (1) (see Fig. 5-4).

| $t$ | $x$ | $y$ |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 1/5 | 1/12 |
| 2 | 1 | 1 |
| 3 | 3 | 83/12 |

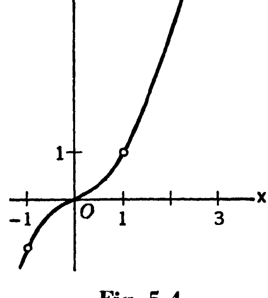



Fig. 5-4

If we want more precise information about the curve we can, for instance, find the first and second derivatives of $y$ with respect to $x$ by differentiation with respect to $t$. The basic principle here is that the derivative of $y$ with respect to $x$ is $dy$ divided by $dx$. Thus we have

$$dx = \frac{1}{10}(1 + 3t^2)\,dt, \qquad dy = \frac{1}{36}(2 + 5t^4)\,dt,$$

$$\frac{dy}{dx} = \frac{10}{36}\frac{2 + 5t^4}{1 + 3t^2} = \frac{5}{18}\frac{2 + 5t^4}{1 + 3t^2}.$$

This shows that the slope of the curve is always positive. In particular, the slope at the origin is $\frac{5}{9}$ (found by putting $t = 0$). For the second derivative we note that

$$\frac{d^2y}{dx^2} = \frac{dy'}{dx}, \quad \text{where} \quad y' = \frac{dy}{dx}.$$

Now
$$dy' = \frac{5}{18}\frac{(1 + 3t^2)20t^3 - (2 + 5t^4)6t}{(1 + 3t^2)^2}\,dt,$$
$$= \frac{5}{9}\frac{15t^5 + 10t^3 - 6t}{(1 + 3t^2)^2}\,dt.$$

Therefore
$$\frac{d^2y}{dx^2} = \frac{50}{9}\frac{15t^5 + 10t^3 - 6t}{(1 + 3t^2)^3}.$$

This expression reveals something which we could not easily discover merely by plotting points, namely, that the curve is concave *downward* for small positive values of $t$. In fact, if we write the numerator of the foregoing expression in the form

$$t(15t^4 + 10t^2 - 6),$$

and solve the biquadratic $15t^4 + 10t^2 - 6 = 0$, we obtain

$$t^2 = \frac{-10 + \sqrt{460}}{30} \sim 0.38, \qquad t \sim 0.62.$$

For $t$ between 0 and this positive root of the biquadratic, the curve is concave downward; for larger values of $t$ the curve is concave upward.

Parametric representation occurs naturally if we think of a curve that is traced out by a moving point. If we establish a time scale, with $t$ the number of time units elapsed after a selected initial instant, the point $(x, y)$ on the curve can be located at various instants by showing how $x$ and $y$ depend on $t$ (i.e., by exhibiting $x$ and $y$ as functions of $t$).

**Example 2:** Let the $x$-axis be horizontal along the ground, and let the

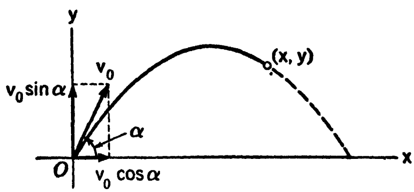



**Fig. 5-5**

$y$-axis be vertical. Let a stone be thrown from the origin, starting up at an angle $\alpha$ with the ground, and with initial speed $v_0$ in this slanting direction (see Fig. 5-5). If air resistance is neglected, the stone will move as follows, according to the laws of mechanics: The $x$-coordinate will increase steadily, with $dx/dt = v_0 \cos \alpha$. The only force acting on the stone after it is thrown is gravity. This causes the $y$-coordinate to change just as though the stone were thrown straight upward with initial speed $v_0 \sin \alpha$. Hence $d^2y/dt^2 = -g$, and $dy/dt = v_0 \sin \alpha$ when $t = 0$. The result is that we have

$$x = v_0 t \cos \alpha, \qquad y = -\frac{1}{2} g t^2 + v_0 t \sin \alpha. \tag{3}$$

Here, then, are equations which represent the path of the stone parametrically.

From the equations (3) we can show that the stone follows a parabolic path. This demonstration is made by expressing $t$ in terms of $x$ from the first equation, and substituting into the second equation:

$$t = \frac{x}{v_0 \cos \alpha}, \qquad y = -\frac{1}{2} g \frac{x^2}{v_0^2 \cos^2 \alpha} + x \tan \alpha. \tag{4}$$

Since $y$ is a quadratic function of $x$, this shows that the point $(x, y)$ moves on a parabola. Further discussion of this example is left for the exercises.

Parametric representation may arise naturally in a geometric way, as the following example shows.

**Example 3:** Suppose $0 < b < a$. Draw two concentric circles of radii $a$, $b$ with centers at the origin. Draw any ray from the origin, cutting the circles at $Q$ and $R$, as in Fig. 5-6. Denote by $\theta$ the angle from the positive $x$-axis to the ray $OQ$. Now let $P$ be the intersection of the line parallel to the $y$-axis through $Q$ and the line parallel to the $x$-axis through $R$. For each angle $\theta$ there is thus determined a point $P$, whose coordinates $(x, y)$ are dependent upon $\theta$. Let it be required to express $x$ and $y$ as functions of $\theta$, and to discuss the curve which is thus represented parametrically.

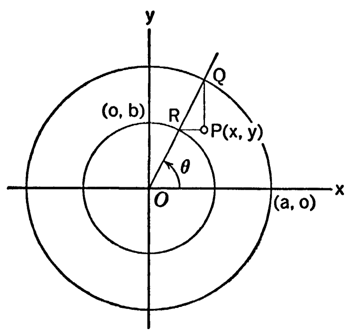



**Fig. 5-6**

Since $OQ = a$ and $OR = b$, we see that

$$x = a \cos \theta, \qquad y = b \sin \theta. \tag{5}$$

These are the required parametric equations. An inspection of the situation shows that as $\theta$ increases from 0 to $\pi/2$, then $P$ goes from $(a, 0)$ to $(0, b)$ along a curve in the first quadrant. This curve is part of an ellipse. In fact,

$$\frac{x}{a} = \cos \theta, \quad \frac{y}{b} = \sin \theta, \quad \text{and so} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

The complete ellipse is traced out as $\theta$ goes from 0 to $2\pi$.

For this case also we illustrate the finding of $y'$ and $y''$ from the parametric representation.

$$dx = -a \sin\theta\, d\theta, \qquad dy = b\cos\theta\, d\theta,$$

$$y' = \frac{dy}{dx} = -\frac{b}{a}\operatorname{ctn}\theta.$$

$$dy' = -\frac{b}{a}(-\csc^2\theta\, d\theta),$$

$$y'' = \frac{dy'}{dx} = -\frac{b}{a^2}\frac{\csc^2\theta}{\sin\theta} = \frac{-b}{a^2\sin^3\theta}.$$

## EXERCISES

1. If $x$ and $y$ are linear functions of the parameter $t$, what kind of a locus is thus represented parametrically? As particular examples consider (a) $x = 2t - 1$, $y = 3t + 4$, and (b) $x = 1 + t$, $y = 1 - t$.

2. Discuss the locus represented parametrically by $x = 1 + t^2$, $y = 2 - t^2$. Describe how $x$ and $y$ vary as $t$ increases from large negative to large positive values. Obtain an equation in rectangular coordinates that is satisfied by all points of the locus. Is every point which satisfies this equation on the locus?

3. Show that the curves represented parametrically by the following pairs of equations are parabolas or parts of parabolas. Indicate which ones are entire parabolas; when only a part of the parabola is represented, indicate which part. In each case describe the way in which the point $(x, y)$ moves as $t$ goes from large negative to large positive values. Also, calculate $dy/dx$ and $d^2y/dx^2$ in each case, from the parametric representation.
   (a) $x = 3t$, $y = \frac{9}{16}t^2$.
   (b) $x = -4 + 4t - t^2$, $y = 4 - 2t$.
   (c) $x = \frac{1}{4}t^2 - 1$, $y = t + 2$.
   (d) $x = t^4$, $y = -t^2$.
   (e) $x = 2\cos\pi t$, $y = 4\sin^2\pi t$.
   (f) $x = \sqrt{1 + t^2}$, $y = 1 + t^2$.

4. Show that the curves represented parametrically by the following pairs of equations are either circles, ellipses, or hyperbolas, or parts of such curves. In each case identify the type of curve, and tell whether all of the curve, or if not, which part of the curve, is represented parametrically. Also calculate $dy/dx$ and $d^2y/dx^2$ in each case, from the parametric representation.
   (a) $x = 3\sin\theta$, $y = 5\cos\theta$.
   (b) $x = \sqrt{2 + t}$, $y = \sqrt{2 - t}$.
   (c) $x = 4\sin\theta$, $y = 4\cos\theta$.
   (d) $x = 2 + 4\cos\phi$, $y = 1 - 2\sin\phi$.
   (e) $x = 1 + 3\cos\theta$, $y = -1 + 3\sin\theta$.

(f) $x = -5\sqrt{\frac{1}{2} + t}$, $y = 3\sqrt{\frac{1}{2} - t}$.
(g) $x = \frac{3}{2}\sqrt{1 + t^6}$, $y = t^3$.
(h) $x = a \sec \theta$; $y = b \tan \theta$.

**5.** Show that the equations

$$x = \frac{2t}{1 + t^2}, \qquad y = \frac{1 - t^2}{1 + t^2}$$

represent the circle $x^2 + y^2 = 1$ except for the point $(0, -1)$. Describe the position of $(x, y)$ on the curve (a) for $t < -1$; (b) for $-1 \leq t \leq 1$; (c) for $t > 1$.

**6.** Show that the equations

$$x = \frac{2t}{1 - t^2}, \qquad y = \frac{1 + t^2}{1 - t^2}$$

represent all of a certain hyperbola except the point $(0, -1)$. Describe the position of $(x, y)$ on the curve (a) for $t < -1$; (b) for $-1 < t < 1$; (c) for $1 < t$.

**7.** Study the curves given by each of the following parametric representations. Think of $t$ as time and examine the way in which $x$ and $y$ vary as functions of $t$, considering all allowable values of $t$, negative as well as positive. Use information obtainable from $dx/dt$ and $dy/dt$, and also from $dy/dx$ and $d^2y/dx^2$, these latter derivatives being expressed as functions of $t$. Draw the curves.

(a) $x = t^2$, $y = (t - 1)^2$.
(b) $x = \dfrac{1}{t + 1}$, $y = t^2$.
(c) $x = t + 2$, $y = 1 + 4/t$.
(d) $x = t^2 - 1$, $y = t^3 - t$.

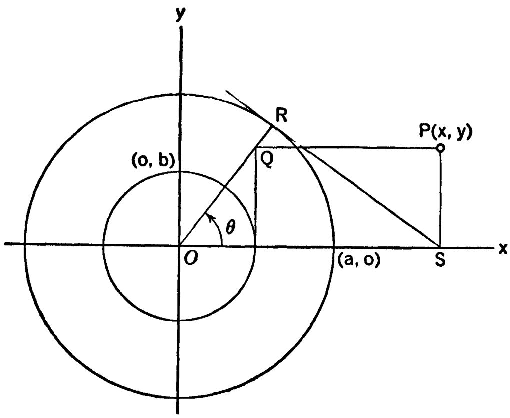



**Fig. 5-7**

**8.** Let $P(x, y)$ be located by the construction indicated in Fig. 5-7. (For some values of $\theta$ the point $Q$ is beyond $R$ on the ray $OR$.) Obtain a parametric representation of the locus followed by $P$ as $\theta$ varies. Show that the locus is the hyperbola $(x^2/a^2) - (y^2/b^2) = 1$.

9. The following questions deal with the motion of the stone, as discussed in Example 2. (a) Locate the vertex of the parabolic path. (b) The *horizontal range* is defined as the distance from 0 to where the stone reaches the $x$-axis on its descent. Show that for fixed $v_0$ and varying $\alpha$ the stone will have the greatest horizontal range when $\alpha = \pi/4$. (d) If $v_0 =$ 100 feet per second, find the two values of $\alpha$ which are required for the thrower of the stone to hit an object $156\frac{1}{4}$ feet away from him on the level. Compare the times of flight of the stone in these two cases. (d) If the ground stretches away from 0 in a straight line of slope $\tan\theta$ (where $-\pi/2 < \theta < \pi/2$), what is the farthest distance along this incline to which the stone can be thrown, assuming $v_0$ is fixed? What value of $\alpha$ will yield this maximum distance?

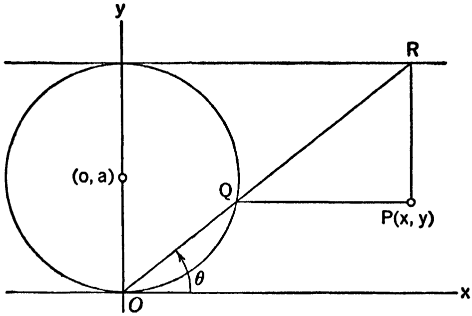



Fig. 5-8

10. Let $P(x, y)$ be located by the construction shown in Fig. 5-8 (where $0 < \theta < \pi$). (a) Obtain parametric equations of the locus followed by $P$ as $\theta$ varies. (b) Find $y'$ and $y''$ (derivatives of $y$ with respect to $x$) in terms of $\theta$. (c) For what values of $\theta$ is $P$ at a point of inflection on the curve? (d) Find an equation of the curve in rectangular coordinates. This curve is called the *witch of Agnesi.*

## 5-8 Cycloids and Other Roulettes

The cycloid is the curve which is traced out by a point on the circumference of a circle when the circle rolls on a straight line in its own plane. A cycloid is most conveniently represented parametrically. Suppose the rolling circle has radius $a$, and let the circle roll on the $x$-axis, starting

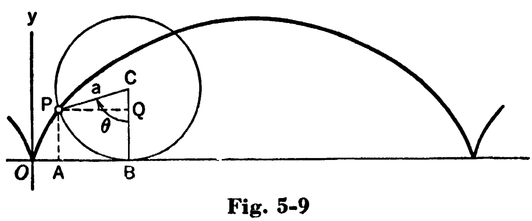



Fig. 5-9

from a position in which the center of the circle is on the positive $y$-axis. We follow the point $P$ on the circle which is at $O$ when the center $C$ of the circle is on the $y$-axis. Let $\theta$ be the angle through which the radius $CP$ has turned when the circle has rolled to a new position. See Fig. 5-9. If $P$ has coordinates $(x, y)$, we see that $y = BC - QC = a - a\cos\theta$. The rolling of the circle implies that $OB = \text{arc } BP = a\theta$. Hence $x = OB - AB = a\theta - a\sin\theta$. Thus the cycloid has the parametric representation

$$x = a(\theta - \sin\theta), \qquad y = a(1 - \cos\theta). \tag{1}$$

These equations are valid for all values of $\theta$, even though they were derived from a diagram in which $\theta$ is a positive acute angle.

From equations (1) we have

$$\frac{dy}{dx} = \frac{a\sin\theta\, d\theta}{a\,d\theta - a\cos\theta\, d\theta} = \frac{\sin\theta}{1 - \cos\theta} = \operatorname{ctn}\frac{\theta}{2}, \tag{2}$$

$$y'' = \frac{d^2y}{dx^2} = \frac{dy'}{dx} = \frac{-\frac{1}{2}\csc^2\frac{\theta}{2}\, d\theta}{a(1 - \cos\theta)\, d\theta} = \frac{-1}{4a\sin^4\frac{\theta}{2}}. \tag{3}$$

Notice that the first and second derivatives are not defined when $\theta = 0$, $\pm 2\pi$, $\pm 4\pi$, etc. These values of $\theta$ correspond to the points where the cycloid meets the $x$-axis; these points are called *cusps*. The tangent to the cycloid becomes parallel to the $y$-axis at the cusps. In between cusps the curve is concave downward.

There are many aspects of the cycloid which are interesting in connection with mechanical problems.

**Example 1:** Prove that the tangent to the cycloid at $P$ (in Fig. 5-9) passes through the top of the rolling circle.

The top of the circle has coordinates $(a\theta, 2a)$. The slope of the tangent at $P$ is given by (2). Hence the equation of the tangent at $P$ is

$$y - a(1 - \cos\theta) = \frac{\sin\theta}{1 - \cos\theta}(x - a\theta + a\sin\theta).$$

We substitute $x = a\theta$ in this equation, and solve for $y$; the result is

$$y = a - a\cos\theta + \frac{\sin\theta}{1 - \cos\theta}\cdot a\sin\theta,$$

$$y = \frac{a(1 - \cos\theta)^2 + a\sin^2\theta}{1 - \cos\theta} = 2a.$$

Hence the tangent at $P$ does indeed go through the point $(a\theta, 2a)$.

When one curve $C$ rolls without slipping on another curve $C'$ whose position is fixed, the locus traced out by a point $P$ which stays fixed on the moving curve $C$ is called a *roulette*. The cycloid is a particular roulette. Other interesting roulettes can be generated by rolling one circle on another.

If a circle rolls on the *inside* of a fixed circle (both circles in the same plane), the locus of a point on the rolling circle is called a *hypocycloid.* If a circle rolls on the *outside* of a fixed circle, the locus of a point on the rolling circle is called an *epicycloid.* We shall show how to represent a hypocycloid parametrically. Let radii of the fixed and rolling circles be $a$, $b$, respectively. Let the fixed circle have its center at the origin and let the initial point of tangency of the circles be at $A$ on the positive $x$-axis, and let us follow the point $P$ which was initially at $A$ (see Fig. 5-10). With $\theta$ and $\phi$ as indicated

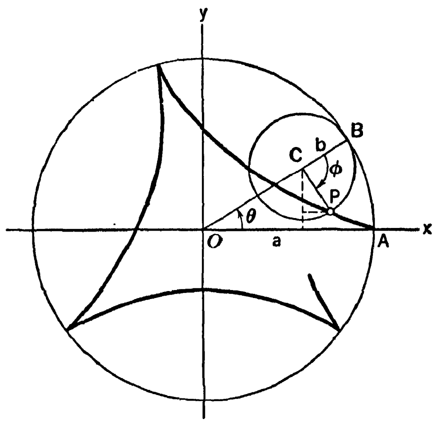



Fig. 5-10

in the diagram, the condition of rolling is expressed by the equality of the arcs $AB$ and $BP$: $a\theta = b\phi$. The coordinates of $P$ are easily seen to be

$$x = (a - b) \cos \theta + b \cos (\phi - \theta),$$
$$y = (a - b) \sin \theta - b \sin (\phi - \theta).$$

But $\phi - \theta = \dfrac{a - b}{b}\theta$. Hence we have

$$x = (a - b) \cos \theta + b \cos \frac{a - b}{b}\theta,$$
$$y = (a - b) \sin \theta - b \sin \frac{a - b}{b}\theta. \tag{4}$$

The arc length along the fixed circle between successive cusps of the hypocycloid is $2\pi b$. If $a/b$ is an integer $n$, the hypocycloid will have $n$ cusps, and the point $P$ will return to $A$ after the smaller circle has rolled off its circumference $n$ times on the fixed circle. We leave it for the student to consider when $P$ will return to $A$ if $a/b$ is a rational fraction, but not an integer, e.g., $a/b = \frac{7}{2}$. What is the situation if $a/b$ is irrational?

The parametric representation of a hypocycloid of four cusps can be put in an especially simple form by using some trigonometric identities. Let $a = 4b$. Then (4) become

$$x = 3b \cos \theta + b \cos 3\theta, \qquad y = 3b \sin \theta - b \sin 3\theta.$$

Now

$$\cos 3\theta = \cos (2\theta + \theta) = \cos 2\theta \cos \theta - \sin 2\theta \sin \theta$$

$$= (2 \cos^2 \theta - 1) \cos \theta - 2 \sin^2 \theta \cos \theta = [2 \cos^2 \theta - 1 - 2(1 - \cos^2 \theta)] \cos \theta$$
$$= 4 \cos^3 \theta - 3 \cos \theta.$$

In the same way we find that

$$\sin 3\theta = 3 \sin \theta - 4 \sin^3 \theta.$$

Hence our parametric equations become

$$x = 4b \cos^3 \theta = a \cos^3 \theta, \qquad y = 4b \sin^3 \theta = a \sin^3 \theta. \tag{5}$$

From this representation it is easy to pass to an equation in rectangular coordinates, namely

$$x^{2/3} + y^{2/3} = a^{2/3}.$$

The curve is shown in Fig. 5-11.

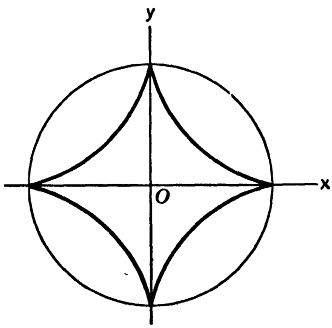



**Fig. 5-11**

**Example 2:** Consider the tangent to the four-cusped hypocycloid at a point in the first quadrant. Show that the length of the part of this tangent cut off by the coordinate axes is always the same, namely, $a$.

To find the slope of the tangent, we have

$$\frac{dy}{dx} = \frac{3a \sin^2 \theta \cos \theta \, d\theta}{-3a \cos^2 \theta \sin \theta \, d\theta} = -\tan \theta.$$

The equation of the tangent is

$$y - a \sin^3 \theta = -\tan \theta \, (x - a \cos^3 \theta).$$

The intercept of this tangent on the $x$-axis is found by setting $y = 0$ and solving for $x$:

$$x = a \cos^3 \theta + a \sin^2 \theta \cos \theta = a \cos \theta.$$

The $y$-intercept is found in a similar manner. It is $y = a \sin \theta$. Hence the length of the portion of the tangent in the first quadrant is

$$\sqrt{a^2 \cos^2 \theta + a^2 \sin^2 \theta} = a.$$

## EXERCISES

1. Show directly, from the equation of the normal to the cycloid at $P$ (in Fig. 5-9), that the normal passes through the point $B$.

2. (a) Supposing that the circle in Fig. 5-9 rolls at a constant rate, with the center $C$ moving $c$ units per second, find the rates of change of the coordinates of $P$. (b) What is the greatest rate of increase of $y$, and for what value of $\theta$ is it attained? (c) What is the greatest rate of increase of $x$, and where is $P$ when this is attained?

3. Show that the area included between the $x$-axis and one arch of the cycloid in Fig. 5-9 is three times the area of the rolling circle. Suggestion: If $S$ is the area $OAP$, we know that

$$\frac{dS}{dx} = y, \quad \text{and hence} \quad \frac{dS}{d\theta} = \frac{dS}{dx}\frac{dx}{d\theta} = y\frac{dx}{d\theta}.$$

It is then possible to compute $S$ as a function of $\theta$ by antidifferentiation. To get the complete area, what value of $\theta$ is wanted? For a helpful clue at one stage of the work, see § 5-5.

4. Show that the hypocycloid for which $a = 2b$ is just the diameter of the fixed circle along the $x$-axis. If in this case the center $C$ of the rolling circle travels with constant angular velocity $\omega$, show that $P$ shuttles back and forth on the $x$-axis with simple harmonic motion, its maximum speed being $a\omega$.

5. (a) If the four-cusped hypocycloid is generated by the small circle rolling so that $d\theta/dt = \omega$ (a constant), find the position of $P$ in the first quadrant when $y$ is increasing most rapidly. (b) What is this greatest rate of increase of $y$?

6. Construct the tangent at the point $P(x, y)$ of the four-cusped hypocycloid (5). Let $Q$ be the foot of the perpendicular from $O$ to this tangent. Find the coordinates of $Q$ in terms of $\theta$. Let $M$ be the mid-point of $PQ$. Show that the locus of $M$ is the circle $x^2 + y^2 = a^2/4$.

7. Derive parametric equations for an epicycloid generated by rolling a circle of radius $b$ externally on a fixed circle of radius $a$. Use a diagram analogous to Fig. 5-10. Draw the epicycloids corresponding to the cases $a = 4b$, $a = 2b$, and $a = b$, respectively.

8. Find the maximum and minimum values of $x$ and $y$, (a) on the epicycloid of Exercise 7 for which $a = b$; (b) on the epicycloid for which $a = 2b$; (c) on the epicycloid for which $a = 4b$.

**9.** For the epicycloid with $a = b$ (as in Exercise 7) show that the tangent at the point corresponding to the parameter value $\theta$ is perpendicular to the tangent at the point corresponding to the parameter value $\theta + \pi$. Show, moreover, that these two tangents intersect at the point $(-3a \cos 2\theta, -3a \sin 2\theta)$, and hence that the locus of their point of intersection is the circle $x^2 + y^2 = 9a^2$.

*CHAPTER VI*

# THE DEFINITE INTEGRAL

## 6-1 The Integral Concept

There are two fundamental concepts which underlie the whole of calculus. These concepts are: *the derivative of a function*, and *the integral of a function.* We are now going to begin a systematic study of the second of these two concepts.

The notion of the definite integral of a function arises by generalization from the idea of finding the area bounded by the lines $x = a$, $x = b$, the $x$-axis, and the graph of $y = f(x)$, where $f$ is a function which is continuous and such that $f(x) \geq 0$ when $a \leq x \leq b$. The idea of finding such an area by a limiting process, using the areas of rectangles, has been explained in § 2-6 (see especially the discussion relating to Figs. 2-22, 2-23, 2-24, and 2-25). The student should review § 2-6 at this time.

Let $f$ be any function which is continuous when $a \leq x \leq b$. The interval from $a$ to $b$, inclusive, is denoted by $[a, b]$. We no longer require that $f(x) \geq 0$. If $n$ is a positive integer and if points $x_0, x_1, \cdots, x_n$ are chosen so that $a = x_0 < x_1 < x_2 < \cdots < x_n = b$, we consider sums formed in the following manner: Choose points $t_1, \cdots, t_n$ in such a way that $x_0 \leq t_1 \leq x_1$, $x_1 \leq t_2 \leq x_2$, and so on. Then consider the sum

$$J = f(t_1)(x_1 - x_0) + f(t_2)(x_2 - x_1) + \cdots + f(t_n)(x_n - x_{n-1}). \qquad (1)$$

We regard the function $f$ and the interval $[a, b]$ as fixed, but the integer $n$ and the points $x_i$, $t_i$ may be chosen in various ways. In most cases the number $J$ obtained in this way will vary as we vary $n$ and the choice of the $x_i$'s and $t_i$'s. However, if we increase $n$, and space the points $x_0, x_1, \cdots, x_n$

in such a way that the maximum of the distances between consecutive points approaches 0 as $n \to \infty$, it is a fact, and a very important one, that the values of $J$ approach a certain limiting value $I$. This limiting value is called *the definite integral of f from a to b*. For the present we shall denote this value $I$ by $I_a^b(f)$ to indicate the part played by $f$ and the interval $[a, b]$ in arriving at the value of $I$.

It is convenient to write $\Delta x_i = x_i - x_{i-1}$. Then, by definition,

$$I_a^b(f) = \lim \, [f(t_1)\,\Delta x_1 + \cdots + f(t_n)\,\Delta x_n]. \qquad (2)$$

The meaning of "limit" here is the following: The absolute difference $|J - I_a^b(f)|$ approaches 0 as the maximum of the values $\Delta x_1, \cdots, \Delta x_n$ is made to approach 0.

In seeking to appreciate the integral concept, let us recall that there were two aspects of our early acquaintance with the derivative concept. We first met the concept in particular applications of it to things like velocity, acceleration, and slope. From these particular cases we passed to the general concept of $f'(x)$ as the limit of a certain quotient. The situation is similar, but more complicated, with the definite integral. There are various possible interpretations of $I_a^b(f)$, and a good deal of our work in calculus has to do with applications of the definite integral in geometry and physics. At the same time, in order to develop enough theory to be able to work problems easily, we must learn to think of the definite integral without being obliged to think of any *particular* interpretation of the integral.

### *The Integral and Areas*

Consider the graph $y = f(x)$, $a \le x \le b$. We shall give an interpretation of the number $J$ expressed as a sum in (1), and of the limit $I_a^b(f)$ as expressed in (2). For illustration see Fig. 6-1 and Fig. 6-2. If $f(t_i) > 0$, then $f(t_i)\,\Delta x_i$ is the area of a certain rectangle *above* the $x$-axis. If $f(t_i) < 0$

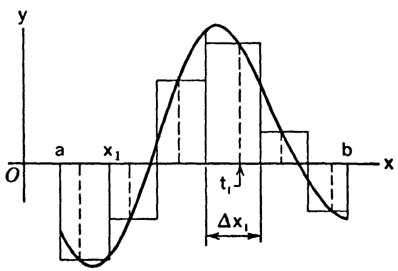



**Fig. 6-1**

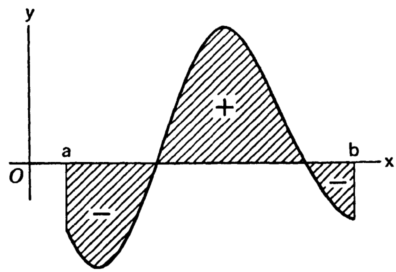



**Fig. 6-2**

(as with $f(t_1)$ in Fig. 6-1), then $f(t_i)\,\Delta x_i$ is the *negative* of the area of a rectangle below the $x$-axis. The number $J$ is then the *algebraic* sum of the areas of $n$ rectangles, with the area of a rectangle counted positively or negatively according as the rectangle is above or below the $x$-axis. The limiting value of $J$, which is the definite integral $I_a^b(f)$, is the algebraic sum of the areas between the graph of $y = f(x)$ and the $x$-axis from $x = a$ to $x = b$, with areas above the $x$-axis counted positively, and those below the $x$-axis counted negatively.

Figures 2-24 and 2-25 give us additional insight into the interpretation of the sums $J$ for various choices of the points $t_1, \cdots, t_n$. These figures illustrate the case in which $f(x) \geq 0$ when $a \leq x \leq b$. One possibility is to choose $t_i$ so that $f(t_i)$ is the smallest value of $f(x)$ for $x_{i-1} \leq x \leq x_i$. Let this smallest value be denoted by $m_i$. Then for $J$ we have what is called a *lower sum*, denoted by $\underline{J}$:

$$\underline{J} = m_1 \Delta x_1 + m_2 \Delta x_2 + \cdots + m_n \Delta x_n. \tag{3}$$

In Fig. 2-24 the value of this lower sum is represented as the sum of the areas of the shaded rectangles. Another possibility is to choose $t_i$ so that $f(t_i)$ is the largest value $M_i$ of $f(x)$ for $x_{i-1} \leq x \leq x_i$. The corresponding $J$ is called an *upper sum* and denoted by $\bar{J}$:

$$\bar{J} = M_1 \Delta x_1 + M_2 \Delta x_2 + \cdots + M_n \Delta x_n. \tag{4}$$

In Fig. 2-25 the value of $\bar{J}$ is represented by the sum of the areas of the shaded rectangles. No matter how $t_i$ is chosen, we have $m_i \leq f(t_i) \leq M_i$, and hence

$$\underline{J} \leq J \leq \bar{J}. \tag{5}$$

In the limiting process we choose all the subintervals so small that each of the differences $M_i - m_i$ is very small. The possibility of doing this depends upon the fact that $f$ is continuous; an exact analysis of what is involved here depends upon a property called *uniform continuity*, which is a subject for study at a more advanced level. If $\epsilon$ is the maximum value of $M_i - m_i$ for $i = 1, 2, \cdots, n$, we see that

$$\bar{J} - \underline{J} = (M_1 - m_1)\,\Delta x_1 + \cdots + (M_n - m_n)\,\Delta x_n \leq \epsilon(b - a).$$

We can use $\underline{J}$ or $\bar{J}$ or any intermediate value of $J$ as an approximation to the value of the integral $I_a^b(f)$. Since

$$\underline{J} \leq I_a^b(f) \leq \bar{J}, \tag{6}$$

as it is not hard to prove, none of these approximations differs from the value of the integral by more than $\epsilon(b - a)$.

**Example 1:** Find the upper and lower sums in the case of $f(x) = (1 + x^3)^{-1}$ and the interval $[0, 2]$, using 8 equal subintervals.

In this case $f(x)$ decreases as $x$ increases, so that $m_i = f(x_i)$ and $M_i = f(x_{i-1})$. We have $\Delta x_i = \frac{1}{4}$ for each $i$, and $x_0 = 0$, $x_1 = \frac{1}{4}$, $x_2 = \frac{1}{2}$, and so on. We prepare a table of values by computing $f(x)$ in fractional form and converting to a decimal.

| $x$ | $f(x)$ |
|---|---|
| $x_0 = 0$ | $1.000 = y_0$ |
| $x_1 = \frac{1}{4}$ | $0.985 = y_1$ |
| $x_2 = \frac{1}{2}$ | $0.889 = y_2$ |
| $x_3 = \frac{3}{4}$ | $0.703 = y_3$ |
| $x_4 = 1$ | $0.500 = y_4$ |
| $x_5 = \frac{5}{4}$ | $0.339 = y_5$ |
| $x_6 = \frac{3}{2}$ | $0.229 = y_6$ |
| $x_7 = \frac{7}{4}$ | $0.157 = y_7$ |
| $x_8 = 2$ | $0.111 = y_8$ |

Now
$$\underline{J} = \tfrac{1}{4}(y_1 + y_2 + \cdots + y_8),$$
$$\overline{J} = \tfrac{1}{4}(y_0 + y_1 + \cdots + y_7).$$

Computation from the table gives $\underline{J} = 0.978$ and $\overline{J} = 1.20$. Hence

$$0.978 < I_0^2(f) < 1.20.$$

*The Integral and Volumes*

A general discussion of volumes may be made in a manner similar to the discussion of areas in § 2-6. Such a general discussion would entail consideration of the integral concept for functions of two and three independent variables. However, the volumes of certain kinds of solids may be discussed easily in terms of the integral concept for functions of one independent variable. The simplest solids to consider are *solids of revolution.*

Consider a function $f$ which is continuous and such that $f(x) \geq 0$ when $a \leq x \leq b$, and think of the portion of the $xy$-plane bounded by $x = a$,

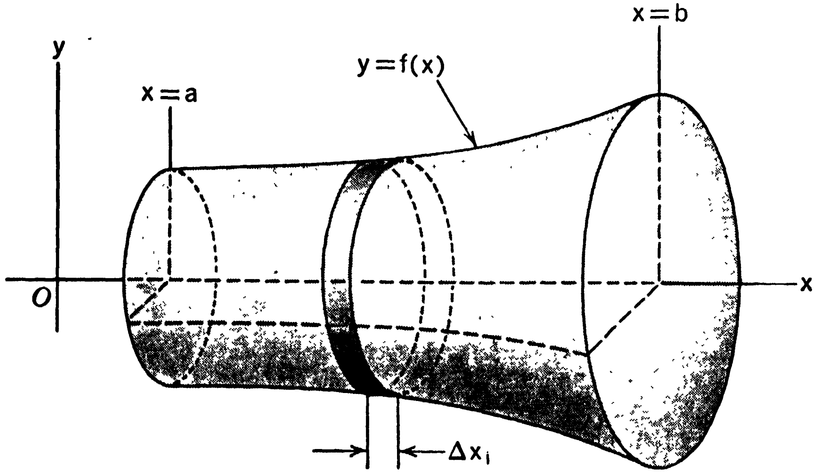



Fig. 6-3

$x = b$, the $x$-axis, and the curve $y = f(x)$. If this plane piece is revolved around the $x$-axis, it generates what is called a solid of revolution (see Fig. 6-3).

The volume of this solid of revolution may be expressed as a definite integral. Let the interval $[a, b]$ be divided into $n$ parts of lengths $\Delta x_1, \cdots, \Delta x_n$, exactly as was done earlier. Then our solid of revolution will be sliced into $n$ circular slabs by the plane sections corresponding to $x = x_0, \cdots, x = x_n$. Consider a typical one of these slabs, of thickness $\Delta x_i$, between $x = x_{i-1}$ and $x = x_i$. Figure 6-4 shows the edge-on view of such a typical slab; as in our previous notation, $m_i$ and $M_i$ denote the smallest and largest values, respectively, of $f(x)$ when $x_{i-1} \leq x \leq x_i$. It is clear that this typical slab is contained in a slice of thickness $\Delta x_i$ cut from a cylinder of radius $M_i$. Hence if $\Delta V_i$ is the volume of our typical slab, then $\Delta V_i \leq \pi M_i^2 \Delta x_i$. Likewise, it is clear that $\pi m_i^2 \Delta x_i \leq \Delta V_i$, for our typical slab contains all of a slice of thickness $\Delta x_i$ cut from a cylinder of radius $m_i$. We see, therefore, that the volume $V$ of our solid of revolution must be not less than

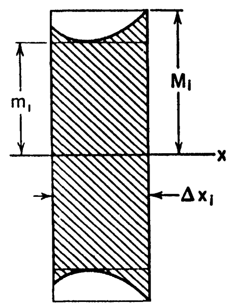



Fig. 6-4

$$\pi(m_1^2 \Delta x_1 + m_2^2 \Delta x_2 + \cdots + m_n^2 \Delta x_n) \tag{7}$$

and not greater than

$$\pi(M_1^2 \Delta x_1 + M_2^2 \Delta x_2 + \cdots + M_n^2 \Delta x_n). \tag{8}$$

But the expressions (7) and (8) are the lower and upper sums, respectively, for the definite integral of $\pi f^2$ from $x = a$ to $x = b$. As the maximum of $\Delta x_1, \cdots, \Delta x_n$ approaches 0, each of these sums approaches the definite integral $I_a^b(\pi f^2)$. Hence this integral is exactly the volume of the solid of revolution:

$$V = I_a^b(\pi f^2). \tag{9}$$

**Example 2:** Apply the foregoing method to find the volume of a right circular cone of altitude $h$ and radius of base $b$.

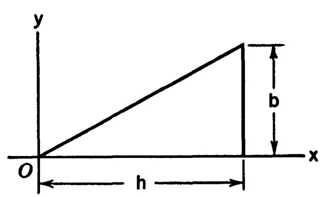



Fig. 6-5

We think of the cone as being generated by revolving about the $x$-axis the triangle in Fig. 6-5. In this case $f(x) = bx/h$ and the interval is $[0, h]$, so we wish to find the integral from 0 to $h$ of $\pi b^2x^2/h^2$. If we divide $[0, h]$ into $n$ equal parts of lengths $h/n$, we see that

$$m_i = f(x_{i-1}) = \frac{b}{h}\frac{i-1}{n}h = \frac{b}{n}(i-1).$$

Hence the sum (7) becomes

$$\frac{\pi b^2}{n^2}[0^2 + 1^2 + \cdots + (n-1)^2]\frac{h}{n}.$$

We use formula (3) in § 2-6 to express this in the form

$$\frac{\pi b^2 h}{n^3}\frac{(n-1)n(2n-1)}{6} = \frac{\pi b^2 h}{6}\left(2 - \frac{3}{n} + \frac{1}{n^2}\right).$$

The limit as $n \to \infty$ is the required volume, so

$$V = \frac{\pi b^2 h}{3}.$$

## EXERCISES

1. Calculate approximating sums for the integral of $x^2$ from $x = 1$ to $x = 9$, as follows: (a) Using 4 equal subintervals and calculating upper and lower sums; (b) using 4 equal subintervals, with $t_1 = 2$, $t_2 = 4$, $t_3 = 6$, $t_4 = 8$; (c) using 8 equal subintervals, and calculating upper and lower sums; (d) using 8 equal subintervals, with $t_i$ midway between $x_{i-1}$ and $x_i$.

2. Calculate an approximating sum for the integral of $\sin x$ from $x = 0$ to $x = \pi/2$, using 2 equal subintervals, $t_1 = \pi/6$, $t_2 = \pi/3$.

3. Calculate 3 approximating sums for the integral of $x^3$ from $x = 0$ to $x = 10$, using 5 equal subintervals: (a) lower sum; (b) upper sum; (c) the sum of the type of formula (1) with $t_i$ midway between $x_{i-1}$ and $x_i$.

4. Calculate an approximating sum for $I_0^8(f)$, where $f(x) = \sqrt{64 - x^2}$, using 4 equal subintervals and $t_1 = 1$, $t_2 = 3$, $t_3 = 5$, $t_4 = 7$.

5. Compute the upper and lower approximating sums for $I_{-1}^2(f)$, where $f(x) = \frac{1}{4}x^2 + 1$, using 12 equal subintervals. Draw a figure and mark the rectangles corresponding to upper and lower sums.

6. Suppose $0 < a < b$. Find the value of $I_a^b(f)$, where $f(x) = x^3$. Use lower sums, with $x_0 = a$, $x_1 = ar$, $x_2 = ar^2$, $\cdots$, $x) = ar^n$, where $r = (b/a)^{1/n}$, and find the limit of the lower sums as $n \to \infty$. Note that $r \to 1$ as $n \to \infty$. In simplifying the expression for the lower sum use the formula for the sum of a geometric progression.

7. The area between the parabola $y^2 = 8x$ and the line $x = 2$ is revolved about the $x$-axis, thus generating a solid of revolution. (a) Compute the upper and lower sums (7) and (8) for this case, using 4 equal subintervals. (b) Find the exact volume of the solid of revolution by finding the limit of the upper sums, using $n$ equal subintervals and letting $n \to \infty$.

**8.** The semicircular area on the right of the $y$-axis and inside the circle $x^2 + y^2 = a^2$ is revolved about the $x$-axis, thus generating a hemisphere. (a) Compute the upper and lower sums (7) and (8) for this case, using 8 equal subintervals. (b) Find the exact volume of the hemisphere by finding the limit of upper sums, using $n$ equal subintervals.

## 6-2 Properties of the Definite Integral

The definition of $I_a^b(f)$ was given in formula (2) of § 6-1. We now begin a process of developing rules and theorems about integrals, for the purpose of arriving at a practical method of calculating the values of definite integrals. Although the definition itself furnishes a direct method of approximating the value of $I_a^b(f)$ to whatever degree of accuracy may be desired, it is possible to develop rules by which the values of many integrals can be found without the need for considering approximating sums. The culmination of this development is found in Theorem 6-D of § 6-4.

Suppose that $a < b < c$, and that $f$ is continuous on the whole interval $[a, c]$. Then we can consider the three integrals $I_a^b(f)$, $I_b^c(f)$, and $I_a^c(f)$. It is an important fact that

$$I_a^c(f) = I_a^b(f) + I_b^c(f). \tag{1}$$

If we interpret the integral in terms of areas, formula (1) represents the fact that the area expressed by $I_a^c(f)$ is the algebraic sum of the areas expressed by $I_a^b(f)$ and $I_b^c(f)$ (see Fig. 6-6). The truth of (1) in general, without recourse to geometrical interpretation, can be traced back to the definition of the integral as the limit of a sum. We use $b$ as one of the points $x_i$ in the subdivision of the interval $[a, c]$. Then, when we form the approximating sum for $I_a^c(f)$, part of the sum is an approximating sum for $I_a^b(f)$, and the rest is an approximating sum for $I_b^c(f)$.

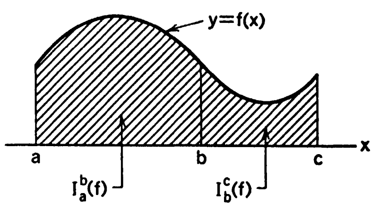


**Fig. 6-6**

It is convenient to be able to use the notation $I_a^b(f)$ even when $a \geq b$. If $a > b$, we shall define $I_a^b(f)$ to be $-I_b^a(f)$. Also, we define $I_a^a(f)$ to be 0. In view of these notational agreements, (1) turns out to be true regardless of the relative positions of $a$, $b$, $c$ on the number scale.

Next, suppose that the functions $f$ and $g$ are each continuous on the interval $[a, b]$. Then

$$I_a^b(f + g) = I_a^b(f) + I_a^b(g). \tag{2}$$

This is seen to be true by examining the definition of $I_a^b(f + g)$. Since the value of $f + g$ at $t_i$ is just $f(t_i) + g(t_i)$, an approximating sum for $I_a^b(f + g)$ is an approximating sum for $I_a^b(f)$ plus an approximating sum for $I_a^b(g)$.

Formula (2) then follows when we take limits. In this argument, as well as in the justification of (1), we depend upon a principle which is a generalization of Theorem 1-C (in § 1-8). Roughly stated, this principle asserts that the limit of the sum of two variable things is the sum of their limits. For the present we shall not attempt a more formally correct statement of this principle.

Another useful formula is the following, in which $c$ denotes any constant factor:

$$I_a^b(cf) = cI_a^b(f). \tag{3}$$

The foregoing rules help us in much the same way as we are helped by knowing the rules for sums and constant factors in connection with differentiation (Theorems 3-A and 3-B in § 3-2).

## 6-3 The Mean-Value Theorem

Before coming to the main subject of this section, we must mention an important fact about continuous functions; this fact is used in the subsequent logical development of this section (in the proof of Theorem 6-B) as well as many places elsewhere later on in the development of calculus.

THEOREM 6-A. *Suppose that $a < b$ and that $f$ is continuous for each value of $x$ such that $a \le x \le b$. Suppose also that $f(a) \ne f(b)$. Then, if $k$ is any number between $f(a)$ and $f(b)$, there is at least one $x$ between $a$ and $b$ for which $f(x) = k$.*

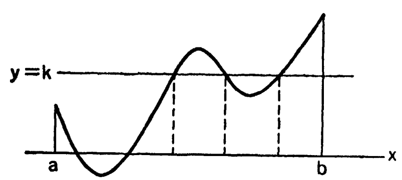


**Fig. 6-7**

This is often called the *intermediate-value theorem* for continuous functions. It is proved in books on advanced calculus. It is this theorem which justifies the representation of the graph of $y = f(x)$ for $a \le x \le b$ as an unbroken curve. Every line $y = k$ for which $k$ is between $f(a)$ and $f(b)$ is crossed at least once by the graph of $y = f(x)$. See Fig. 6-7, in which three such crossings are shown.

The following theorem is of crucial importance in our development of information about definite integrals.

THEOREM 6-B (*Mean-value theorem*). *If $a < b$ and if $f$ is continuous on $[a, b]$, then there is some $X$ such that $a \le X \le b$ and*

$$I_a^b(f) = (b - a)f(X). \tag{4}$$

*Proof.* Let $m$ and $M$ denote the minimum and maximum values of $f$ on $[a, b]$. Then, if $J$ is any approximating sum for $I_a^b(f)$ (see (1) in § 6-1), it is readily seen that

$$m(b - a) \leq J \leq M(b - a),$$

for $m \, \Delta x_i \leq f(t_i) \, \Delta x_i \leq M \, \Delta x_i$, and the sum of all the $\Delta x_i$'s is $(b - a)$. When we pass to the limit we obtain the inequalities

$$m(b - a) \leq I_a^b(f) \leq M(b - a). \tag{5}$$

Now let

$$\mu = \frac{1}{b - a} I_a^b(f). \tag{6}$$

This number $\mu$ is called the *arithmetic mean* of the values of $f(x)$ on $[a, b]$. We see from (5) and (6) that $m \leq \mu \leq M$. In order to arrive at (4), all that now remains is to show that there is some $X$ such that $\mu = f(X)$. Then (4) is a consequence of (6). Now $m$ and $M$ are values of $f$ at certain points $x_1$ and $x_2$ on $[a, b]$; we know this from Theorem 2-A, § 2-1. Hence, by Theorem 6-A, each number between $m$ and $M$ must be attained as a value of $f$ at some point between $x_1$ and $x_2$. Since $m \leq \mu \leq M$, it follows that there is some $X$, either between $x_1$ and $x_2$ or coinciding with one of them, such that $f(X) = \mu$. This finishes the argument.

## 6-4 The Fundamental Relations Between Derivatives and Integrals

In this section we shall learn about the important connections between differentiation and integration. Differentiation is the process of passing from a function to its derivative. Integration is the process of passing from a function to a definite integral of it.

To start our investigation we adopt a slightly different but vitally important point of view about definite integrals. Instead of thinking of $I_a^b(f)$ where $a$ and $b$ are fixed, we think of $I_a^x(f)$, *where $x$ is variable.* Then $I_a^x(f)$ depends on $x$, and hence defines a function of $x$. Our first result concerns the derivative of this function.

THEOREM 6-C. *Suppose $a$ and $x$ are points of an interval on which $f$ is continuous. Keeping $a$ fixed, but regarding $x$ as variable, let us define*

$$G(x) = I_a^x(f). \tag{1}$$

*Then $G$ has a derivative given by the formula*

$$G'(x) = f(x). \tag{2}$$

*Proof.* We know that, by definition,

$$G'(x) = \lim_{h \to 0} \frac{G(x + h) - G(x)}{h}. \tag{3}$$

Now, by (1),

$$G(x + h) - G(x) = I_a^{x+h}(f) - I_a^x(f).$$

We know from (1) in § 6-2 that

$$I_a^{x+h}(f) = I_a^x(f) + I_x^{x+h}(f).$$

Hence
$$\frac{G(x+h) - G(x)}{h} = \frac{1}{h} I_x^{x+h}(f).$$

Now we use Theorem 6-B, which permits us to write

$$I_x^{x+h}(f) = hf(X),$$

$X$ being some number between $x$ and $x + h$. Thus

$$\frac{G(x+h) - G(x)}{h} = f(X). \tag{4}$$

Letting $h$ approach 0, we see that $X$ must approach $x$; since $f$ is continuous, this implies that $f(X) \to f(x)$. Consequently, from (3) and (4) we conclude that $G'(x) = f(x)$, as asserted in (2).

If the student will now reread § 2-7 as far as (3), he will see that this early part of § 2-7 can be regarded as a geometrical interpretation of the proof of Theorem 6-C, with the integral interpreted as an area.

Theorem 6-C exhibits one half of the fundamental relation between differentiation and integration. The other half is exhibited in the next theorem.

THEOREM 6-D. *Suppose that $f$ is continuous on $[a, b]$, and suppose that in some way we are able to find a function $F$ which has a derivative such that $F'(x) = f(x)$ for each $x$ on $[a, b]$. Then we can find the value of the integral $I_a^b(f)$ by the formula*

$$I_a^b(f) = F(b) - F(a). \tag{5}$$

*Proof.* Consider the relation between the function $F$ here described and the function $G$ defined by (1) in the preceding theorem. Since $F'(x) = G'(x) = f(x)$, we see that

$$\frac{d}{dx}[F(x) - G(x)] = 0$$

for each $x$ on $[a, b]$. It then follows from one of the fundamental items (item V) in § 2-1 that $F(x) - G(x)$ has the same value at all points of $[a, b]$. In particular then, $F(a) - G(a) = F(b) - G(b)$, or

$$F(b) - F(a) = G(b) - G(a). \tag{6}$$

But by the definition of $G$, we see that

$$G(b) - G(a) = I_a^b(f) - I_a^a(f). \tag{7}$$

Since $I_a^a(f) = 0$, we see that (6) and (7) combine to give us the desired relation (5).

If $f$ is a given function and $F$ is another function such that $F'(x) = f(x)$, $F$ is called an *antiderivative* of $f$; this terminology was first used in § 2-2. Finding $F$ when $f$ is given is called *antidifferentiation.* Theorem 6-D shows that integration can be accomplished by antidifferentiation if the necessary antiderivative can be found. This theorem justifies us in introducing a new notation for the definite integral, based on the standard notation for antiderivatives, as introduced in § 5-3. Hereafter we shall mainly use

$$\int_a^b f(x)\,dx \quad \text{in place of} \quad I_a^b(f).$$

The numbers $a$, $b$ affixed to the symbol $\int$ are called the *limits of integration.* The function $f$ is called the *integrand.* The symbol $\int$ is now called "the integral sign."

With this new notation for a definite integral it is immaterial what letter is used to denote the independent variable of the integrand. That is,

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(u)\,du,$$

and so on. This is apparent from the very definition of $I_a^b(f)$; it is also apparent from (5), for $F(b) - F(a)$ is the same, no matter whether we write $F(x)$, $F(t)$, or $F(u)$ for a typical value of $F$.

We now illustrate the technique based on Theorem 6-D. The following useful notations will be employed:

$$\Big[F(x)\Big]_a^b = F(x)\Big|_a^b = F(b) - F(a).$$

**Example 1:** Find the value of

$$\int_{-1}^{2} (12x^5 - 16x^3 + 9x^2 - 2)\,dx.$$

An antiderivative of

$$12x^5 - 16x^3 + 9x^2 - 2 \text{ is } 2x^6 - 4x^4 + 3x^3 - 2x.$$

Therefore

$$\begin{aligned}\int_{-1}^{2} (12x^5 - 16x^3 + 9x^2 - 2)\,dx &= \left(2x^6 - 4x^4 + 3x^3 - 2x\right)\Big|_{-1}^{2}\\ &= [2(64) - 4(16) + 3(8) - 4] - [2 - 4 - 3 + 2]\\ &= 84 - (-3) = 87.\end{aligned}$$

In order to be adept in the evaluation of definite integrals, the student must develop skill in finding antiderivatives. For the present, it will suffice to review the standard formulas in § 5-3 and § 5-5, and to review the simple substitution techniques in § 5-4.

**Example 2:** Find the value of $\displaystyle\int_{-4}^{4\sqrt{3}} \frac{dx}{16 + x^2}.$

We know that

$$\int \frac{dx}{16 + x^2} = \frac{1}{4}\tan^{-1}\frac{x}{4} + C.$$

Hence

$$\int_{-4}^{4\sqrt{3}} \frac{dx}{16 + x^2} = \frac{1}{4}\tan^{-1}\frac{x}{4}\Big|_{-4}^{4\sqrt{3}} = \frac{1}{4}\tan^{-1}(\sqrt{3}) - \frac{1}{4}\tan^{-1}(-1).$$

The student must remember that in all the standard calculus formulas involving inverse trigonometric functions, *principal values* are used. For review see § 4-4. In the present case $\tan^{-1}\sqrt{3} = \pi/3$, $\tan^{-1}(-1) = -\pi/4$. Thus

$$\int_{-4}^{4\sqrt{3}} \frac{dx}{16 + x^2} = \frac{\pi}{12} + \frac{\pi}{16} = \frac{7\pi}{48}.$$

## EXERCISES

**1.** Find the value of each definite integral.

(a) $\int_{-2}^{4} (8 - 2x + x^2)\,dx.$

(b) $\int_0^{2b} \left(b^2u - \frac{1}{2}u^3 + \frac{u^5}{16b^2}\right) du.$

(c) $\int_8^{27} (-2x^{-2/3} + 12x^{1/3} - 6x^{2/3})\,dx.$

(d) $\int_0^1 \sqrt{4 - 3x}\,dx.$

(e) $\int_1^2 \frac{dx}{(3x - 2)^{5/2}}.$

(f) $\int_0^{\sqrt{3}} \frac{x\,dx}{(4 - x^2)^{3/2}}.$

(g) $\int_{-1}^{4} \frac{u\,du}{(u^2 + 9)^2}.$

(h) $\int_0^2 (1 + t^3)^{1/2}\,t^2\,dt.$

(i) $\int_0^{\pi/5} \sin 5\theta\,d\theta.$

(j) $\int_{\pi/2}^{2\pi/3} \sec^2 \frac{x}{2}\,dx.$

(k) $\int_{7\pi/12}^{11\pi/12} \frac{\cos 2\theta}{\sin^2 2\theta}\,d\theta.$

(l) $\int_{-\sqrt{3}}^{1} \frac{dx}{1 + x^2}.$

(m) $\int_0^1 \frac{dx}{\sqrt{4 - x^2}}.$

(n) $\int_{-3/2}^{\sqrt{3}/2} \frac{dx}{9 + 4x^2}.$

(o) $\int_{-1}^{2} \sqrt{4 - x^2}\,dx.$

(p) $\int_0^1 \frac{u\,du}{1 + u^4}.$

**2.** Find the value of each definite integral.

(a) $\int_{-b}^{b} (b^4 - 2b^2t^2 + t^4)\,dt.$

(b) $\int_0^a (a^{2/3} - x^{2/3})^3\,dx.$

(c) $\int_{-1/3}^{2/3} \frac{dx}{\sqrt{2 + 3x}}.$

(d) $\int_0^{2a} \frac{x\,dx}{\sqrt{a^2 + x^2}}.$

(e) $\int_{2b}^{3b} \frac{x\,dx}{(x^2 - b^2)^3}.$

(f) $\int_0^1 \frac{x^3}{(1 + x^4)^2}\,dx.$

(g) $\int_{-\pi/6}^{2\pi/3} \cos 3\theta\,d\theta.$

(h) $\int_0^{2\pi} \sin \frac{\pi - x}{3}\,dx.$

(i) $\int_0^{\pi/6} \csc^2\left(\frac{\pi}{3} - y\right) dy.$

(j) $\int_0^{\pi/9} \frac{\sin 3x}{\cos^2 3x}\,dx.$

(k) $\int_{-1/3}^{0} \frac{dt}{1 + 9t^2}.$

(l) $\int_{2/3}^{2/\sqrt{3}} \frac{dx}{\sqrt{16 - 9x^2}}.$

(m) $\int_{-1/\sqrt{2}}^{1} \frac{du}{\sqrt{1 - (u^2/2)}}.$

(n) $\int_{5/4}^{5\sqrt{3}/4} \frac{dx}{25 + 16x^2}.$

(o) $\int_{3/5}^{6/5} \sqrt{36 - 25x^2}\,dx.$

(p) $\int_0^{\sqrt{2}} \frac{x\,dx}{\sqrt{16 - x^4}}.$

**3.** Complete the following statements.

(a) If $F(x) = \int_0^x \sin t\, dt$, then $F'(x) =$

(b) The derivative with respect to $y$ of $\int_4^y \sqrt{x}\, dx$ is $\cdots$.

(c) $\dfrac{d}{du}\displaystyle\int_0^u \sqrt{1 + t^4}\, dt =$ (e) $\dfrac{d}{dx}\displaystyle\int_x^{\pi/2} \cos\theta\, d\theta =$

(d) $\dfrac{d}{ds}\displaystyle\int_1^s \tan^{-1} y\, dy =$ (f) $\dfrac{d}{dx}\displaystyle\int_{-x}^{x} \frac{du}{1 + u^4} =$

**4.** Let $\mu$ be the arithmetic mean of the continuous function $f$ on $[a, b]$. If $a = x_0 < x_1 < \cdots < x_n = b$, where the points $x_0, x_1, x_2, \cdots, x_n$ are equally spaced, explain why

$$\frac{f(x_0) + \cdots + f(x_{n-1})}{n} \quad \text{and} \quad \frac{f(x_1) + \cdots + f(x_n)}{n}$$

both approach $\mu$ as $n \to \infty$.

**5.** Find the arithmetic mean $\mu$ and a value of $X$ such that $f(X) = \mu$ for each of the following cases.

(a) $f(x) = x^2$ on $[0, 2\sqrt{3}]$. (d) $f(x) = \sin x$ on $[0, \pi]$.

(b) $f(x) = x^3$ on $[0, 4]$. (e) $f(x) = \dfrac{1}{(1 + x^2)}$ on $[0, 1]$.

(c) $f(x) = \sqrt{x}$ on $[0, 4]$. (f) $f(x) = (4 - x^2)^{-1/2}$ on $[0, 1]$.

## 6-5 More About Areas

We now consider a region of the following sort: It is the part of the $xy$-plane lying between two curves $y = f_1(x)$ and $y = f_2(x)$, and between the lines $x = a$ and $x = b$, where $f_1$ and $f_2$ are continuous on $[a, b]$. We assume that the curves do not intersect, except possibly at one or both ends of $[a, b]$. The situation is shown in Fig. 6-8. We divide the interval $[a, b]$ into $n$ subintervals by points $x_0, x_1, \cdots, x_n$ and let $\Delta x_i = x_i - x_{i-1}$, just as in § 6-1. Then the several lines $x = x_i$ divide the region whose area we wish to find into $n$ parallel strips of widths $\Delta x_1, \cdots, \Delta x_n$. A single strip resembles a long narrow rectangle; the ends of the strip are curves, however. If $t_i$ is a number such that $x_{i-1} \le t_i \le x_i$, the value of

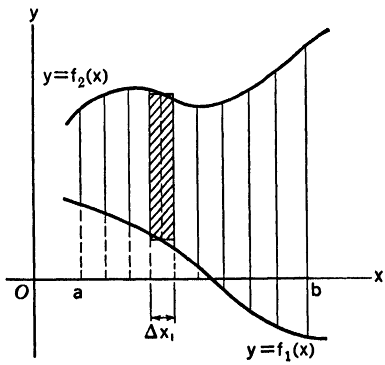



**Fig. 6-8**

$$[f_2(t_i) - f_1(t_i)]\, \Delta x_i \tag{1}$$

is the area of a rectangle which resembles the strip of which we are speaking (the rectangle is shaded in Fig. 6-8). Hence it is reasonable to expect that in the limit as max $(\Delta x_1, \cdots, \Delta x_n) \to 0$, the sum of all the expressions (1) will be the exact area of the region we are considering. Since this limit is the integral of $f_2 - f_1$ from $x = a$ to $x = b$, we obtain the formula

$$A = \int_a^b [f_2(x) - f_1(x)]\, dx \tag{2}$$

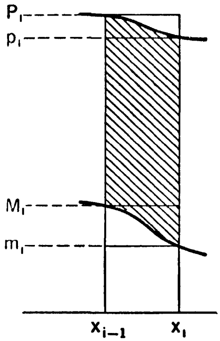


**Fig. 6-9**

for the required area.

In order to be sure that this procedure will give the required area, we can reason as follows. Let $m_i$, $M_i$ denote the minimum and maximum values of $f_1(x)$ on $[x_{i-1}, x_i]$; likewise let $p_i$, $P_i$ denote the minimum and maximum values of $f_2(x)$ on the same subinterval. Then the area $\Delta A_i$ of the $i$th strip of our region certainly satisfies the following inequalities (see Fig. 6-9, in which $\Delta A_i$ is shaded).

$$(p_i - M_i)\,\Delta x_i \le \Delta A_i \le (P_i - m_i)\,\Delta x_i.$$

But the limit of the sum of the terms $(p_i - M_i)\,\Delta x_i$ is the same as the limit of

$$(p_1\,\Delta x_1 + \cdots + p_n\,\Delta x_n) - (M_1\,\Delta x_1 + \cdots + M_n\,\Delta x_n),$$

and we know this limit to be

$$\int_a^b f_2(x)\,dx - \int_a^b f_1(x)\,dx = \int_a^b [f_2(x) - f_1(x)]\,dx.$$

The same kind of argument applies to the sum of the terms $(P_i - m_i)\,\Delta x_i$. Hence (2) is correct, by the basic principle set forth in § 2-6.

The student should not memorize formula (2) and apply it in mere routine fashion to the exercises following this section. It is the idea of expressing the area as the limit of a sum of areas of rectangles which is the important thing. The method applies equally well to finding the area of a region by dividing it into strips parallel to the $x$-axis. In this case the width of a typical strip will be $\Delta y$, and the area will be found by integration with respect to $y$.

As an aid to the student we give an outline of the steps to be followed in finding an area by integration.

(i) Draw a figure showing the region whose area is to be found. Mark the coordinate axes plainly. Write down the equations of the curves and lines which form the boundary of the region.

(ii) Decide upon a method of dividing the region into thin parallel strips and draw a typical strip on the figure.

(iii) Using the equations of the curves which bound the region, calculate the length of a typical strip (in terms of $x$ if the width is $\Delta x$, in terms of $y$ if the width is $\Delta y$), and write down the expression for the area of the rectangle which serves as an approximation to the area of the strip.

(iv) Set up the integral which is the limit of the sum of the expressions of which a typical one was found in step (iii). Observe that if the expression for the area of the rectangle is $F(x)\,\Delta x$, the integral will be

$$\int_a^b F(x)\,dx.$$

The limits of integration are found by examining the figure.

(v) Carry out the integration and find the value of the definite integral.

We give an illustrative example in which the strips are taken parallel to the $x$-axis.

**Example:** Find the area enclosed between the parabolas $y^2 = -4(x - 1)$ and $y^2 = -2(x - 2)$.

The curves and a typical strip parallel to the $x$-axis are shown in Fig. 6-10.

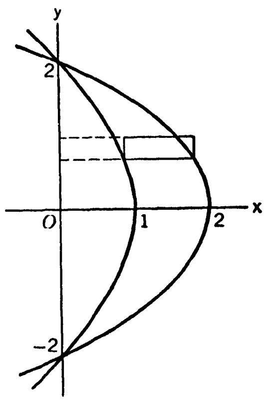



**Fig. 6-10**

Both parabolas are symmetric with respect to the $x$-axis, and they open to the left. We solve each equation for $x$ in terms of $y$:

$$x = \frac{1}{4}(4 - y^2), \qquad x = \frac{1}{2}(4 - y^2).$$

The difference of these two values of $x$ gives the length of a typical rectangle. Hence the area of the rectangle is $\frac{1}{4}(4 - y^2)\,\Delta y$. The area is therefore

$$A = \int_{-2}^{2} \frac{1}{4}(4 - y^2)\,dy.$$

Note the limits of integration. It is evident from symmetry that we may integrate from 0 to 2 and double the result. Thus

$$A = \frac{1}{2}\int_0^2 (4 - y^2)\,dy = \frac{1}{2}\left[4y - \frac{1}{3}y^3\right]_0^2 = \frac{8}{3}.$$

## EXERCISES

**1.** In each part of the exercise an area is described; express the area as a definite integral and compute its value. Solve (b), (c), (g), and (j) in two ways: once integrating with respect to $x$, and once with respect to $y$.
(a) Between $y = 9 - x^2$ and $y = x^2$.
(b) Between $y = x^3$, the $y$-axis, and $y = -27$.
(c) Between $4y = x^3$ and $y = x$, $x \geq 0$.
(d) Between one arch of the curve $y = 3 \cos 2x$ and the $x$-axis.
(e) Between $y = \sin x$ and $y = -3 \sin x$, $\pi/3 \leq x \leq \pi$.
(f) Between $y = 9 - x^2$ and $(x + 3)^2 = -4y$.
(g) Between $y^2 = 4x^3$ and $y^2 = 16x$, $y \leq 0$.
(h) Between $y^2 = -16(x - 1)$ and $3y^2 = 16(x + 3)$.
(i) Between $9x = y^3$ and $y^2 = -3(x - 6)$, $-3 \leq y \leq 3$.
(j) Between $4y = x^3$ and the tangent to this curve at $x = -2$.

**2.** Proceed as directed in Exercise 1. Solve (d), (f), and (j) in two ways.
(a) Between $y^2 = 4(x - 2)$ and $5y^2 = 8(x + 4)$.
(b) Between $2y^2 + 9x = 36$ and $3x + 2y = 0$.
(c) Between $y = x^3 - 4x$ and $y = 5x$, $x \geq 0$.
(d) Bounded by $2y = (x - 1)^3$, $y = 4$, and $x = -2$.
(e) Bounded by $2x + 3y + 1 = 0$ and $x + 3 = (y - 1)^2$.
(f) Bounded by $4x = 8 + y^3$, $x = -14$, and $y = 4$.
(g) Between $y = x^4 - 4x^2$ and $y = -4$.
(h) Between $y = x^4 - 4x^2$ and the semicircle $y = \sqrt{4 - x^2}$.
(i) Between $y(4 + x^2) = 16$ and $y = 2$.
(j) Between $27y = -2x^3$ and the tangent to this curve at $x = 3$.

**3.** Find the area:
(a) Between the curve $y^3(x - 1)^2 = 1$ and the $x$-axis, $2 \leq x \leq 9$.
(b) Between the curve $y(7 - 4x)^3 = 5$ and the $x$-axis, $-1 \leq x \leq 0$.
(c) Between the upper part of the parabola $y^2 = 3 + 3x$ and the $x$-axis, $-1 \leq x \leq 11$.

**4.** Find the area:
(a) Between the curve $y\sqrt{4 - x^2} = 8$ and the $x$-axis, $-1 \leq x \leq \sqrt{2}$.
(b) Between the curve $y(2x^2 + 1) = 3$ and the $x$-axis, $-1/\sqrt{2} \leq x \leq \sqrt{6}/2$.

**5.** Two lines are drawn from the origin, tangent to the circle $(x - 5)^2 + y^2 = 5$. Find the area enclosed between the two lines and the circle.

**6.** Find the area bounded by $x^2y = 1$, $y = -27x$, and $-8y = x$.

**7.** Find the area of the three-sided figure between the parabolic arc $x^{1/2} + y^{1/2} = a^{1/2}$ and the coordinate axes.

**8.** Find the entire area enclosed by the curve $y^2 = 9x^2 - x^4$.

**9.** Consider one of the arches of $y = 6 \sin (x/2)$ above the $x$-axis. Find the area under this arch and above the line $y = 3$.

## 6-6 Three-Dimensional Figures

In this section explanations will be given of the use of rectangular coordinates in space of three dimensions. The primary purpose is to teach the student a few simple things about the description of certain curves and surfaces by equations involving the coordinates $x$, $y$, $z$. Also, there is some illustration of a useful method of drawing diagrams to represent simple solids and surfaces.

A rectangular coordinate system is based on three mutually perpendicular straight lines with a common point of intersection $O$, called the *origin* of coordinates. The lines are called coordinate axes and each plane determined by two of the axes is called a coordinate plane. A positive direction is assigned along each axis, and each axis is provided with a number scale whose zero point is at $O$. We shall assume that the same unit of distance is used on each axis. Figure 6-11 shows how the coordinates of a point $P$ are determined. The positive axes are lettered $x$, $y$, $z$, respectively. The $x$-coordinate of $P$ is defined as the *directed distance* from the $yz$-plane to the point $P$. This distance is measured along the perpendicular to the plane, and is positive or negative according as $P$ is on the same side of the $yz$-plane as the positive or negative portion of the $x$-axis. In Fig. 6-11 the coordinates of $P$ are the directed distances $OA = x$, $AB = y$, $BP = z$.

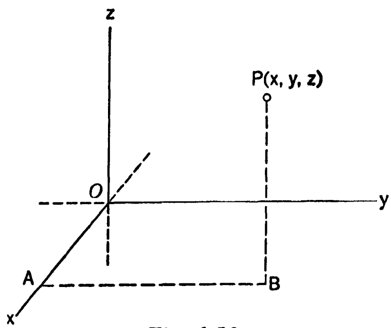


**Fig. 6-11**

The coordinate system shown in Fig. 6-11 is called *right-handed.* If the labels on the positive $x$ and $y$ axes were exchanged it would be called *left-handed.* We shall always use right-handed systems.

The three coordinate planes divide space into eight regions called *octants*. The one in which all the coordinates are positive is called the *first* octant.

To find the distance $\overline{OP}$ in Fig. 6-11, we use the theorem of Pythagoras twice:

$$\overline{OB}^2 = \overline{OA}^2 + \overline{AB}^2 = x^2 + y^2, \qquad \overline{OP}^2 = \overline{OB}^2 + \overline{BP}^2,$$

and so

$$\overline{OP}^2 = x^2 + y^2 + z^2. \tag{1}$$

From this we see that $P$ lies on a sphere of center $O$ and radius $r$ if and only if

$$x^2 + y^2 + z^2 = r^2. \tag{2}$$

Hence (2) is an equation of this sphere. For many problems relating to such a sphere it is convenient to make a diagram representing merely the first-octant portion of the sphere. Figure 6-12 is such a diagram; on it are

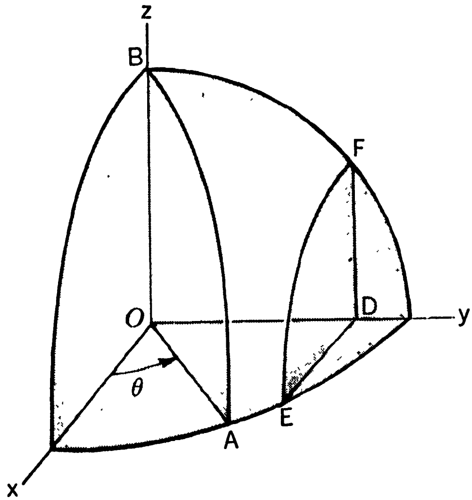



**Fig. 6-12**

shown the sections of this octant of the sphere by certain planes. The section $OAB$ is made by a plane which passes through the $z$-axis and makes an acute angle $\theta$ with the $xz$-plane. The arc $AB$ is, of course, a quarter circle of radius $r$. The section $DEF$ is made by a plane perpendicular to the $y$-axis. The arc $EF$ is a quarter circle. If $OD = a$, where $0 < a < r$, then $DE = DF = \sqrt{r^2 - a^2}$.

In § 6-7 there is explained a method of finding the volume of a solid by integration. One of the essential steps in the method requires that we be able to find the areas of all sections of the solid made by planes perpendicular to some selected line. When the sections are elementary figures such as triangles, rectangles, or circles, the finding of the areas is often a fairly simple matter.

**Example 1:** A solid has the following shape: its base is a circle of radius 2, and plane sections of the solid perpendicular to a fixed diameter $BB'$ of the base are isosceles triangles having chords of the circle as their bases. The third vertex of each isosceles triangle lies along one of the lines $BC$, $B'C$, where $C$ is a point 3 units directly above the center $O$ of the circular base.

To visualize the solid, let the circular base be placed in the $xy$-plane with the center at the origin and the fixed diameter $BB'$ along the $x$-axis. Let the point $C$ fall on the positive $z$-axis. In Fig. 6-13 we show only a quarter of the solid and half of a typical section of the solid by a plane perpendicular to the

$x$-axis. To find the area of the triangle $DEF$, we seek to express the lengths $ED$ and $EF$ in terms of the distance $x = OE$ of the plane section from the origin. Now it is clear that the points on the circle $ADB$ (of radius 2) satisfy the equations

$$z = 0, \qquad x^2 + y^2 = 4. \tag{3}$$

Hence, if $OE = x$, the length $ED$ is the value of $y$ found by solving the second equation in (3). Thus $ED = y = \sqrt{4 - x^2}$. In similar fashion we find $EF$. The line $BC$ is described by the equations

$$y = 0, \qquad \frac{x}{2} + \frac{z}{3} = 1.$$

Solving for $z$ in the second equation, we find $EF = z = \frac{3}{2}(2 - x)$. The area of the triangle $DEF$ is now readily expressed in terms of $x$; it is

$$\frac{1}{2}(DE)(EF) = \frac{1}{2}\sqrt{4 - x^2} \cdot \frac{3}{2}(2 - x) = \frac{3}{4}(2 - x)\sqrt{4 - x^2}.$$

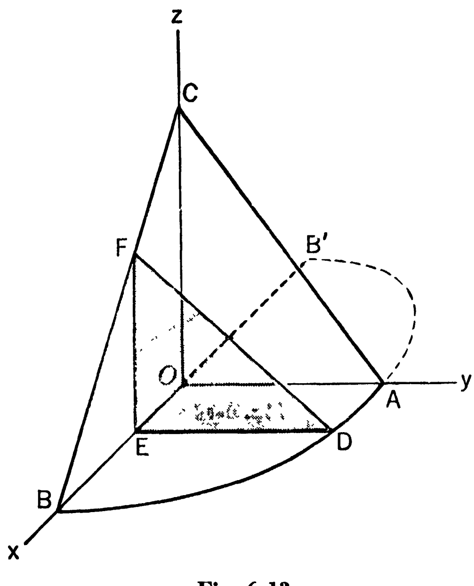



**Fig. 6-13**

We saw in the foregoing Example 1 that the pair of equations (3) describe a circle of radius 2 in the $xy$-plane, with center at the origin. If we omit the equation $z = 0$ and consider the single equation $x^2 + y^2 = 4$, we see readily that this equation is satisfied by those and only those points lying on the circular cylinder of radius 2 with axis along the $z$-axis. It is convenient to consider the word *cylinder* in a more general sense. A surface is called a cylinder if there is a direction in space such that when a line is drawn parallel to this direction and through a point of the surface, all points of the line belong to the surface. Such lines in the surface are called *elements* of the cylinder. In general, a single equation in just two of the three coordinates $x$, $y$, $z$ is the equation of a cylinder. If $z$ (for example) is the missing letter, the elements of the cylinder are parallel to the $z$-axis. In that case the shape of the cylinder is revealed by considering the equation in $x$ and $y$ as the equation of a curve in the $xy$-plane. All sections of the cylinder by planes parallel to the $xy$-plane are exactly congruent curves. For instance, the equation $x^2 = 4y$ represents a parabolic cylinder with elements parallel to the $z$-axis. The cylinder is symmetric with respect to the $yz$-plane. See Fig. 6-14, in which is shown a part of the cylinder cut off between the $xy$-plane and a parallel plane.

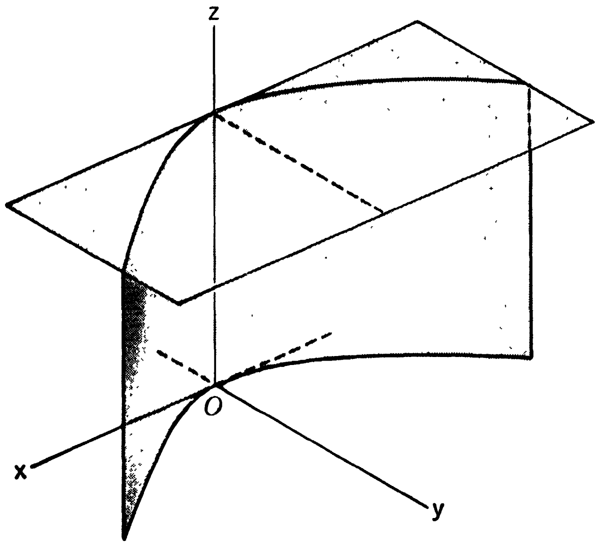



**Fig. 6-14**

**Example 2:** Consider Fig. 6-15. The curve $AED$ is supposed to be a parabola with vertex at $D$ and axis along $DO$. If $OA = 1$ and $OD = 3$, this parabola is described by the equations

$$y = 0, \qquad 3x^2 = -(z - 3). \tag{4}$$

The curve $CGD$ is also supposed to be part of a parabola with vertex at $D$ and axis along $DO$. If $OC = 2$, this parabola is described by the equations

$$x = 0, \qquad 3y^2 = -4(z - 3). \tag{5}$$

The equation $3x^2 = -(z - 3)$ by itself represents a parabolic cylinder. Lines $AB$ and $EF$ are segments of elements of this cylinder. Likewise, the equation $3y^2 = -4(z - 3)$ by itself represents a parabolic cylinder, and lines $BC$ and $FG$ are segments of elements of it. These two cylinders intersect in the first octant along the curve $DFB$. The figure $EFGH$ is a rectangle in a plane parallel to the $xy$-plane. If $OH = z$, the dimensions of this rectangle can be found from (4) and (5). Using (4) we have

$$HE = x = \sqrt{(3 - z)/3},$$

and using (5) we have

$$HG = y = 2\sqrt{(3 - z)/3}.$$

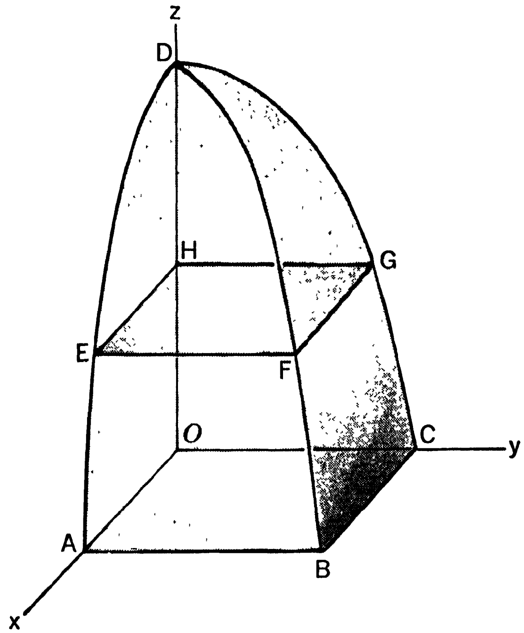



**Fig. 6-15**

References to Fig. 6-13 and Fig. 6-15 will be made in § 6-7. For a fuller discussion of many topics from analytic geometry of three dimensions, the student may refer to Chapter XVIII.

## 6-7 Volumes by Slicing

In § 6-1 it was explained how the volume of a solid of revolution can be expressed as a definite integral. In order to achieve this, the solid is cut into thin slices by a series of planes perpendicular to the axis of symmetry of the solid of revolution. This process may be applied to solids of other types. The essential matter for success of the method is that we be able to find the area of each section of the solid made by a plane perpendicular to some fixed line.

Let us consider planes perpendicular to the $x$-axis, and suppose the solid in which we are interested extends from the plane $x = a$ to the plane $x = b$, where $a < b$. Let $A(x)$ denote the area of the cross section of the solid made by the plane determined by an arbitrary value of $x$. If Fig. 6-13 in § 6-6 is taken as an example, with the triangle $DEF$ as this typical section, then

$$A(x) = \frac{3}{4}(2 - x)\sqrt{4 - x^2}.$$

This was shown in § 6-6. Let us think once more of the general case. We are supposing the solid is sufficiently smooth in shape to make the area $A(x)$ a continuous function of $x$. Now let the interval $[a, b]$ be divided into $n$ parts by points $x_0, \cdots, x_n$, the lengths of the parts being $\Delta x_1, \cdots, \Delta x_n$. The planes perpendicular to the $x$-axis at $x_0, \cdots, x_n$ divide our solid into $n$ slices, and $\Delta x_1, \cdots, \Delta x_n$ are the respective thicknesses of these slices. If $\Delta x_1$ is small, it is plausible to think of $A(x_1)\,\Delta x_1$ as an approximation to the volume of this first slice, because $A(x)$ will not be very much different from $A(x_1)$ when $x_0 \le x \le x_1$. Similar remarks apply to the other slices, and so it is plausible to think of

$$A(x_1)\,\Delta x_1 + \cdots + A(x_n)\,\Delta x_n$$

as an approximation to the actual volume of the solid; moreover, it is plausible to suppose that this sum approaches the exact volume when the maximum of $\Delta x_1, \cdots, \Delta x_n$ approaches zero. This supposition leads to the formula

$$V = \int_a^b A(x)\,dx \tag{1}$$

for the volume of the solid.

Formula (9) in § 6-1 is a special case of the present formula, for in Fig. 6-3 $A(x)$ is the area of a circle of radius $f(x)$, so that $A(x) = \pi[f(x)]^2$.

Our plausibility argument leading to formula (1) cannot be regarded as a valid general *proof* that the volume is given by the formula. Nevertheless, the formula is correct; a genuinely general justification of it would lead us far beyond our main concern of the moment. Further remarks about the general discussion of volumes will be found in the treatment of double and triple integrals, later in the book.

In many special cases it is easy to see more clearly that (1) must be correct. For instance, if we are considering the first octant solid shown in Fig. 6-13, the interval $[a, b]$ is $[0, 2]$. As $x$ increases, both dimensions of the triangle $DEF$ decrease. In this case therefore, the slice of the solid between the planes $x = x_0$ and $x = x_1$ clearly has a volume $\Delta V$ such that

$$A(x_1)\,\Delta x_1 < \Delta V < A(x_0)\,\Delta x_1. \tag{2}$$

Similar remarks apply to the other slices, and so the exact volume of the solid lies between

$$A(x_1)\,\Delta x_1 + \cdots + A(x_n)\,\Delta x_n \qquad \text{(too small)}$$

and

$$A(x_0)\,\Delta x_1 + \cdots + A(x_{n-1})\,\Delta x_n \qquad \text{(too large).}$$

Since both of these sums approach

$$\int_0^2 A(x)\,dx = \int_0^2 \frac{3}{4}(2 - x)\sqrt{4 - x^2}\,dx, \tag{3}$$

this integral gives the exact volume.

**Example:** Find the volume of the first octant solid depicted in Fig. 6-15 (Example 2, § 6-6).

Here the sectioning planes are perpendicular to the $z$-axis. The typical section is a rectangle of area $A(z) = (HE)(HG) = \frac{2}{3}(3 - z)$. Since $OD = 3$, the required volume is

$$V = \int_0^3 \frac{2}{3}(3 - z)\,dz = \frac{2}{3}\left(3z - \frac{1}{2}z^2\right)\Big|_0^3 = 3.$$

## EXERCISES

**1.** The horizontal cross section of a certain pyramid $x$ feet from its top is a square of side $\frac{1}{4}x$ feet. The pyramid is 40 feet high. Find its volume.

**2.** A World's Fair tower is 80 feet tall. A horizontal cross section of the tower $x$ feet from its top is a square of side $\frac{1}{400}(x + 40)^2$ feet. Find the volume of the tower.

**3.** A horn is generated by a circle which moves in the following way: the plane of the circle is perpendicular to the $x$-axis; the center of the circle is in the $xy$-plane, and its diameter in that plane is a line segment cut off between the curves $y = x^{1/3}$ and $3y = x^{1/3}$. Find the volume generated by the circle as its center moves from $x = 0$ to $x = 8$.

**4.** Find the volume of a solid whose base is a circle of radius 5 if all the plane sections perpendicular to a fixed diameter of this base are equilateral triangles.

**5.** Find the volumes of the following solids of revolution:

(a) A sphere of radius $a$;

(b) The spherical segment obtained by rotating about the $x$-axis the area inside the circle $x^2 + y^2 = a^2$ and on the right of the line $x = a - h$, where $0 < h < 2a$;

(c) The *prolate* spheroid obtained by revolving the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ about its major axis $(a > b)$;

(d) The *oblate* spheroid obtained by revolving the ellipse in (c) about its minor axis;

(e) The solid generated by revolving about the $x$-axis the area between $5y^2 = 32x$ and $x = 10$;

(f) The solid generated by revolving about the $y$-axis the area between $x^2 = 8y$ and $y = 2$;

(g) The solid generated by revolving about the $x$-axis the area bounded by $xy = 1$, $x = 1$, $x = 3$, and $y = 0$;
(h) The solid generated by revolving about the $y$-axis the area between the two branches of the hyperbola $9x^2 - 16y^2 = 144$ and between the lines $y = 0$, $y = 3$;
(i) The solid generated by revolving about the $y$-axis the area between the coordinate axes and the curve $x^{1/2} + y^{1/2} = a^{1/2}$.

**6.** Draw the parabola $Hx^2 = B^2y$, where $B$ and $H$ are positive constants.
(a) Find the volume generated when the area between the parabola and the line $y = H$ is revolved about the $y$-axis. Compare your answer with the volume of a cylinder of height $H$ and radius of base $B$.
(b) Find the volume generated when the area between the parabola and the $x$-axis, $0 \leq x \leq B$, is revolved about the $x$-axis.

**7.** A solid has as its base the area bounded by the hyperbola $16x^2 - 9y^2 = 144$ and the line $x = 6$. Every cross section of this solid perpendicular to the $x$-axis is (a) a square, or (b) an equilateral triangle. Find the volume in each case.

**8.** Find the volume of the first octant solid of Fig. 6-13, as described in Example 1, § 6-6.

**9.** Refer to Fig. 6-13 and assume that $ADB$ is a quarter circle. Find the volume of the solid $OABC$ if $OA = a$ and $OC = c$.

**10.** In Fig. 6-13 let $ADB$ be one quarter of an ellipse, with $OA = a$, $OB = b$, $OC = c$. Find the volume $OABC$.

**11.** A solid has as its base the triangle cut from the first quadrant by the line $3x + 4y = 12$. Every plane section of the solid perpendicular to the $y$-axis is a semicircle. Find the volume of the solid.

**12.** Find the volume generated by revolving about the $y$-axis the area between $xy = 4$ and $x + y = 5$.

**13.** The axes of two right circular cylinders, each of radius $a$, intersect at right angles. Find the volume of the space which is inside of both cylinders.

**14.** In felling a tree a woodsman first saws halfway through at right angles to the trunk. He then makes a second cut in a plane inclined at an angle $\theta$ to the first cut, the two planes meeting in a line which intersects the central axis of the tree. Find the volume of the wedge removed if the tree is assumed to be a cylinder of radius $b$.

**15.** A square hole of side 2 inches is cut through a cylindrical post of radius 2 inches. If the axis of the hole intersect the axis of the post at right angles, find the volume cut out (a) assuming that a pair of opposite plane sides of the hole are perpendicular to the axis of the post; (b) assuming that the plane sides of the hole make 45° angles with the axis of the post.

## 6-8 Work

An important concept in physics is that of *work done by a force*. There is a relation between work and energy, as we shall see in § 6-9. Our first concern here is to make clear the *definition* of the work concept; this definition is made in terms of a definite integral.

Consider a moving particle and a force which acts on the particle while it moves. In defining the work done by this force we can ignore other forces which may be acting on the particle at the same time. We suppose that the particle moves from $x = a$ to $x = b$ on the $x$-axis; either $a < b$ or $a > b$ is permitted, but we shall suppose that the particle does not reverse the direction of its motion, and hence that it passes just once through each point between $a$ and $b$. The force we are considering may be variable in magnitude and direction; we shall denote by $f(x)$ the component of the force in the positive $x$-direction when the particle is at $x$ on the axis. This component may be negative. The work done by the force is then *defined* to be

$$W = \int_a^b f(x)\,dx. \tag{1}$$

The unit of work takes its name from the unit of force and the unit of distance. For example, if force is in pounds and distance is in feet, the unit of work is the *foot-pound*. In the CGS system the unit of force is one *dyne*; a dyne-centimeter of work is called an *erg*. In the MKS system the unit of force is one *newton*; a newton-meter of work is called a *joule*. For conversion, 1 newton is $10^5$ dynes, and 1 joule is $10^7$ ergs.

In the case of a constant force directed along the line of motion, the definition of work in (1) gives the result: *work = force times distance*, as used in the study of elementary physics before calculus is available.

According to the definition of work, the component of force at right angles to the direction of motion does no work at all. When a particle *does* reverse the direction of its motion, we split up the motion into parts within which the direction does not change, and add algebraically the contributions to the total work from each part of the motion.

**Example 1:** A ball weighing 4 ounces is thrown up and rises to a height of 60 feet. Find the work done by the force of gravity on the ball from the time the ball is thrown until it is on the way down and 20 feet above its starting point.

The force of gravity on the ball is $\frac{1}{4}$ pound. Hence, if the $x$-axis extends upward, with the origin at the starting point of the ball, $f(x) = -\frac{1}{4}$. The work done while the ball is rising is

$$W_1 = \int_0^{60} -\frac{1}{4}\,dx = -15 \quad \text{foot-pounds},$$

and that done while the ball is falling from $x = 60$ to $x = 20$ is

$$W_2 = \int_{60}^{20} -\frac{1}{4}\,dx = 10 \quad \text{foot-pounds}.$$

Hence the total work done is $W = -15 + 10 = -5$ foot-pounds.

The notion of work occurs in a simple way in connection with problems in extension or compression of elastic materials. When an elastic body, such as a rubber band, a steel wire, or an aluminum bar, is subjected to a pull, it is found by experiment that the body will stretch, and that as long as the applied force is not too great the tension in the body is directly proportional to the amount of the elongation. Thus, in Fig. 6-16, if $OA$ represents an unstretched cord and $OP$ represents the same cord stretched an amount $s$ under a tension $T$, then we have

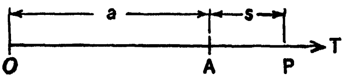


**Fig. 6-16**

$$T = ks, \tag{2}$$

where $k$ is a constant of proportionality.

**Example 2:** An unstretched spring is 3 feet long. When the spring is used to suspend a 25-pound weight it stretches to a length of $4\frac{1}{2}$ feet. Find the work necessary to stretch the spring from a length of $4\frac{1}{2}$ feet to a length of 6 feet.

We begin by finding the proportionality constant $k$ in (2) for this case. Since $T = 25$ pounds when $s = 1\frac{1}{2}$ feet, we have $25 = \frac{3}{2}k$, or $k = \frac{50}{3}$. Thus, in general, $T = \frac{50}{3}s$. To stretch the spring from a length $3 + s_1$ to a length $3 + s_2$ will require work

$$W = \int_{s_1}^{s_2} T\,ds$$

by the force $T$ acting on the point $P$ at the end of the spring. In the present case the work required is found by putting $s_1 = 1\frac{1}{2}$, $s_2 = 3$, and so

$$W = \int_{3/2}^{3} \frac{50}{3}\,s\,ds = \frac{225}{4} = 56.25 \quad \text{foot-pounds}.$$

In formula (2) the constant $k$ depends upon the length of the wire or bar, the area of its cross section, and the material of which it is composed. It has been determined experimentally that for a bar of cross-sectional area $A$ and length $L$ when unstretched we have the formula

$$T = EA\,\frac{s}{L}, \tag{3}$$

where $E$ is a constant of proportionality which depends only on the material of which the bar is made and the units employed for length and force.

The law embodied in (3) is called *Hooke's law* and $E$ is called *Young's modulus*, or the modulus of elasticity.

The situation for compressing a spring or an elastic bar is similar to the situation for stretching. Equation (2) holds with $T$ meaning the compressive force and $s$ the amount of shortening.

## EXERCISES

**1.** If a force of 160 pounds stretches a spring which is naturally 6 feet long to a length of $6\frac{1}{2}$ feet, find (a) the length of the spring when a 240-pound weight is hung up by it; (b) the size of the weight which will stretch the spring to 9 feet; (c) the work done in stretching the spring from 6 to 7 feet, from 7 to 8 feet, and from 8 to 9 feet, respectively; (d) the tension in the spring, and its length, when 4000 foot-pounds of work have been expended in stretching it, starting from the unstretched condition.

**2.** A force of 1000 newtons (the weight of a mass of slightly over 102 kilograms) compresses a spring from its natural length of 1 meter to 0.8 meter. Find (a) the work required to compress the spring from 0.9 to 0.7 meter, and from 0.7 to 0.5 meter, respectively; (b) the work done in compressing the spring to half its natural length; (c) the amount the spring is compressed when half the work done in (b) has been expended (starting with no compression); (d) the ultimate force compressing the spring when 131.25 joules of work have been expended, starting with the spring at a length of 0.9 meter. What is the length of the spring at this ultimate state?

**3.** A particle of mass $M$ grams at the origin attracts a particle of mass $m$ grams at a point $x$ centimeters away on the $x$-axis with a force of magnitude $kmM/x^2$ dynes, where $k$ is the gravitational constant. Find the work done by this force (a) when $m$ moves from $x = 0.01$ to $x = 0.1$; (b) when $m$ moves from $x = 1$ to $x = 0.1$.

**4.** Find the work done by the force of gravitation acting on a satellite of mass $m$ when it is rocketed from the earth's surface to a distance $h$ above that surface. Show that the answer can be put in the form $-\dfrac{mgRh}{R + h}$, where $R$ is the radius of the earth and $g$ is the acceleration of gravity at the earth's surface. See § 5-6, Exercise 5. For $h = 200$ miles how does this result compare with what the work would be if the gravitational force remained constant (equal to its value at the earth's surface)?

**5.** If a particle moves on the $x$-axis under the influence of a force $f(x) = -kx$, where $k$ is constant, the particle moves with simple harmonic motion. The constant $k$ can be evaluated if we know the value $f(x)$ for some particular $x$ different from zero.

(a) If $f(x) = -10$ pounds when $x = 2$ feet, find the work done by the force as $x$ diminishes from 2 to 1.

(b) If $f(x) = 50$ pounds when $x = -\frac{1}{8}$ foot, find the work done by the force as $x$ increases from $-\frac{1}{3}$ to $\frac{1}{4}$.

(c) How much work is done by the force as the particle moves from $x = -b$ to $x = b$, where $b$ is the amplitude of the simple harmonic motion?

**6.** A gas is confined in a cylindrical chamber fitted with a piston. Let the area of cross section of the cylinder be $A$, and let the volume and length of the gas chamber be $V$ and $x$, respectively. If $p$ is the pressure of the gas, the force with which the gas presses on the piston is $f(x) = pA$. Suppose $p$ and $V$ are related by the formula $pV^\gamma = C$, where $\gamma$ and $C$ are constants. This is the relation which holds during an adiabatic expansion or compression of the gas. Show that, if the gas expands from a volume $V_1$ to a volume $V_2$, the work done by the gas pressing on the piston is

$$W = \int_{V_1}^{V_2} p\,dV.$$

Evaluate this for a gas initially occupying 64 cubic inches at a pressure of 128 pounds per square inch, when it expands to 8 times the initial volume; assume $\gamma = \frac{5}{3}$.

**7.** A spring of natural length 3 feet will stretch to 4 feet when a 5-pound weight is suspended in equilibrium at the end of the spring. Suppose the $x$-axis extends downward, with its origin at the upper end of the spring. Let the weight be allowed to oscillate vertically as it hangs on the spring. Find the total work done by gravity and the pull of the spring on the weight as the weight (a) descends from $x = 3$ to $x = 4$; (b) rises from $x = 5$ to $x = 3\frac{1}{2}$.

## 6-9 Energy

The discussion of work and of energy provides a good illustration of the way calculus is used in physics.

Consider a particle of mass $m$ moving along a straight line, which we take to be the $x$-axis. The kinetic energy of the particle is, by definition, $\frac{1}{2}mv^2$, where $v$ is the velocity of the particle ($v = dx/dt$). Let the total force on the particle have a component in the positive $x$-direction of amount $f(x)$ when the particle is at $x$. We shall assume that we confine attention to a part of the motion during which the particle moves entirely in one direction. There is then a theorem which relates the change in the kinetic energy to the work done by the total force acting on the particle.

THEOREM 6-E. *Under the conditions just explained the change in the kinetic energy of the particle, when it traverses a certain interval, is equal to the total work done on it by the force which prevails during the motion.*

*Proof.* Let the particle move from $a$ to $b$ on the $x$-axis. We use Newton's second law in the form

$$mv\frac{dv}{dx} = f(x) \tag{1}$$

(see (2) and (3) in § 5-6). The units are taken so that the proportionality constant is 1 in Newton's law. From (1) we see that

$$\frac{d}{dx}\left(\frac{1}{2}mv^2\right) = f(x);$$

therefore, when we regard $v$ as a function of $x$ and use Theorem 6-D, we see that

$$\frac{1}{2}mv^2 \Big|_a^b = \int_a^b f(x)\,dx. \tag{2}$$

The right side here is the work done by the force; the left side is

$$\left(\frac{1}{2}mv^2\right)_{x=b} - \left(\frac{1}{2}mv^2\right)_{x=a},$$

which is exactly the change in the kinetic energy. Hence (2) expresses the truth of Theorem 6-E.

In certain kinds of physical situations there is value in introducing a concept of potential energy. For instance, if a particle of mass $m$ is lifted up from the ground, we say that we have increased its potential energy, because if we then release the particle and let it fall, the force of gravity does work on the particle and its kinetic energy is increased. When we assign a measure of potential energy to a particle under the influence of a force, we are measuring the capacity of that force to do work on the particle. However, we do not *invariably* associate potential energy with a force; this can be done only for certain kinds of forces. Here we shall mention just a few examples of the potential energy concept.

(1) For a particle of mass $m$ subjected to the constant acceleration of gravity near the surface of the earth, the potential energy $V$ due to the acceleration of gravity is defined as $V = mgy$, where $g$ is the numerical value of the acceleration of gravity and $y$ is the coordinate of the particle. The $y$-axis is taken to have its positive direction upward; the location of the origin is immaterial. Changing the origin will of course change $V$, but this does not matter, for the important thing about potential energy is not its actual value, but the value of $dV/dy$, which controls the change in $V$ for a certain change in $y$. Observe that $-dV/dy = -mg$; this is the force due to gravity on $m$ in the $y$-direction.

(2) For a particle of mass $m$ moving on the positive $x$-axis under the influence of a force $f(x) = -k/x^2$, where $k$ is a constant, the potential energy is defined as $V = -k/x$. Observe that $-dV/dx = -k/x^2 = f(x)$. This is the case of inverse-square-law attraction toward the origin, as in the case of gravitation; see § 5-6, Example 2.

These examples illustrate the point that in appropriate cases the potential energy $V$ due to a force $f(x)$ in the direction of the $x$-axis has the property $-dV/dx = f(x)$. Hence, if the particle moves from $x = a$ to $x = b$, the work and the potential energy are related as follows:

$$W = \int_a^b f(x)\,dx = -\int_a^b \left(\frac{dV}{dx}\right) dx = -V \Big|_a^b. \tag{3}$$

That is, the change in the potential energy from $x = a$ to $x = b$ is the negative of the work done by the force. If we suppose that $f(x)$ is the total force which prevails during the motion, we can combine (2) and (3) to obtain

$$\left(\frac{1}{2}mv^2 + V\right)_{x=b} = \left(\frac{1}{2}mv^2 + V\right)_{x=a}.$$

In other words, the sum of the kinetic and potential energies remains constant as the particle traverses the interval. This is the principle of conservation of energy.

Energy, either kinetic or potential, is measured in the same units as work.

If a bead slides on a rough wire, the force of friction does not have a potential energy associated with it. This shows up in the fact that if the wire extends along the $x$-axis, the friction cannot be represented as a force $f(x)$, for the direction of the frictional force when the particle is at $x$ depends, not merely on $x$, but on the direction in which the particle is moving.

**Example:** Calculate the increase in the potential energy of a 20-pound mass rocketed from the earth's surface to a point 4000 miles up from the surface, and relate it to the kinetic energy and the work done in getting the mass up there.

We shall take the $x$-axis with origin at the center of the earth. Let $R$ (= 4000 miles) be the radius of the earth, $m$ the mass in question, and $F(x)$ the mechanical force applied to the mass by the rocket motor and the drag of what air resistance there is. The gravitational force is $-mgR^2/x^2$ (see § 5-6, Exercise 5). Hence, by Newton's law,

$$mv\frac{dv}{dx} = F(x) - \frac{mgR^2}{x^2}.$$

For an increase of $x$ from $R$ to $2R$ we obtain by integration

$$\frac{1}{2}mv^2\Big|_R^{2R} = \int_R^{2R} F(x)\,dx + mgR^2\left(\frac{1}{2R} - \frac{1}{R}\right),$$

or

$$\frac{1}{2}mv^2\Big|_R^{2R} + \frac{mgR}{2} = \int_R^{2R} F(x)\,dx. \tag{4}$$

The potential energy is $V = -mgR^2/x$, so it changes from $-mgR$ to $-mgR/2$; the increase is $mgR/2$. We see then in (4) that the change in the total energy (kinetic plus potential) is equal to the work done by the applied force $F(x)$. If $R = 4000$ miles and $mg = 20$ pounds, the increase in the potential energy is 40,000 mile-pounds.

### EXERCISES

1. A mass of .1 kilogram is falling freely near the earth's surface. At a certain instant ($t = 0$) the mass is 60 meters above the ground and it is falling 49 meters per second. Take the $y$-axis positively upward, with the origin at ground level. (a) Find the velocity of the mass when it reaches a point 10 meters above ground. (b) Compute the increase of the kinetic energy of the mass between $y = 60$ and $y = 10$. (c) Compute the decrease in the potential energy of the mass between $y = 60$ and $y = 30$. (d) Find the velocity of the mass and its height above the ground at $t = 1$ second. (e) Find the algebraic amounts of the changes in kinetic energy and in potential energy from $t = 0$ to $t = 1$, and the algebraic amount of work done by the force of gravity on the mass during this second.

2. A 30-pound projectile is rocketed upward from the earth. It is cut loose from its rocket motor at a height of 100 miles, its speed then being 5 miles per second. (a) Find the algebraic amount of the work done by gravitation when the projectile rises from 100 miles to 1000 miles. (b) Find the algebraic changes in the kinetic and potential energies during the rise from 500 to 1000 miles above the earth. (c) Find the algebraic change in the potential energy from the time of power cutoff until the projectile stops rising. Take the radius of the earth to be 4000 miles. See § 5-6, Exercise 5.

3. Consider Exercise 12, § 5-6. (a) If a 5-pound particle is dropped into the tunnel at one end, find the change in its kinetic energy by the time it falls to the center of the earth. (b) Find the work done by the gravitational attraction as the particle goes from the surface to a point in the tunnel 2000 miles beyond the center of the earth. (c) If the $x$-axis is taken along the tunnel with origin at the center of the earth, express the gravitational attraction on the particle as a function of $x$, and from this deduce an appropriate definition of the potential energy associated with this force. (d) Use Newton's law to deduce the relation $v^2 = \frac{g}{R}(R^2 - x^2)$, where $v$ is the velocity, and show that this equation is essentially the same as the statement of the principle of conservation of energy for this case.

4. In the theory of elasticity the work done in stretching a bar is regarded as being converted into potential energy stored up in the bar. Show that under a tension $P$ the amount of potential energy stored up in the bar is $P^2L/2AE$. Use formula (3) of § 6-8.

## 6-10 Moments of Inertia

Suppose a rigid body is rotating about a fixed axis. For example, the body might be a sphere spinning about a diameter, or a cube oscillating like a pendulum about a horizontal axis along one of its edges. In the study of such motions it is found to be essential to consider what is called the *moment of inertia* of the body about the axis.

To define moment of inertia we start with the case in which the body is imagined to be a system consisting of a finite number of mass particles rigidly connected by weightless rods. We imagine the system to be rotating about a fixed axis, so that each particle travels in a circle about the axis and the planes of all the circles are perpendicular to the axis. Let the masses be $m_1, \cdots, m_n$, and let the distance from $m_k$ to the axis be $r_k$. Then the moment of inertia of the system about this axis is defined to be

$$I = m_1r_1^2 + \cdots + m_nr_n^2. \tag{1}$$

For rigid bodies in which the mass is distributed continuously (e.g., a solid sphere, a circular disk, a rectangular plate, a hollow cylindrical shell, and so on), the definition of moment of inertia is made by a suitable integral. In order to see how to formulate the definition suitably in such cases, we divide the rigid body into a large number of small pieces and consider the system of particles which we obtain by concentrating all the mass of each piece at some point in the piece. This auxiliary system of particles has a moment of inertia as defined by (1); we then pass to the limit as the number of pieces is increased and their maximum dimension approaches zero. The limit of the moment of inertia of the auxiliary system is then *defined* as the moment of inertia of the rigid body with continuously distributed mass.

The concept of mass-density enters when we consider continuously distributed mass. Suppose, for instance, that mass is distributed along a thin wire or in a thin rod. We can then speak of *linear density* of mass. If the mass of any given length of the rod or wire is a fixed constant times that length, the fixed constant is called the linear density, or mass per unit length. We also consider the case of continuous *variable* linear density. Suppose the rod or wire extends along the $x$-axis from $x = a$ to $x = b$ $(a < b)$. Let $f$ be a continuous function of $x$ on $[a, b]$ such that $f(x) \geq 0$ and such that every subinterval contains at least one point $x$ at which $f(x) > 0$. Then we can conceive of $f$ as a density function, which means that the mass between $x_1$ and $x_2$ (where $x_1 < x_2$) is

$$\int_{x_1}^{x_2} f(x)\, dx,$$

and the entire mass from $a$ to $b$ is

$$M = \int_a^b f(x)\, dx. \tag{2}$$

The average density over the interval $[x_1, x_2]$ is, by definition,

$$\frac{1}{x_2 - x_1} \int_{x_1}^{x_2} f(x)\, dx.$$

As we see from Theorem 6-B, this average density is equal to some value

$f(X)$, where $x_1 \leq X \leq x_2$. If we let the interval close down on a fixed point, the average density approaches as a limit the density at that point.

The numerical value of the density is usually represented by a Greek letter. We shall use the letter $\sigma$. In the foregoing, then, $\sigma = f(x)$.

When we think of mass as spread over a plane area, as in the case of a circular disk, a rectangular plate, and so on, we have a concept of areal density, or mass per unit area. In the case of constant density this is just the quotient: mass divided by area. For the general discussion of variable density in this context we must wait until we discuss double integrals. A similar postponement is also necessary for the concept of variable density in three-dimensional mass distribution.

When the density of a body is constant, we describe the body as being *uniform*, or *homogeneous*.

**Example 1:** We shall show how to formulate the definition of moment of inertia of a straight wire about an axis at right angles to the wire, it being assumed that the axis and the wire are in a single plane. Suppose the wire extends along the $x$-axis from $a$ to $b$, with density $\sigma = f(x)$, and let us find the moment of inertia about the $y$-axis.

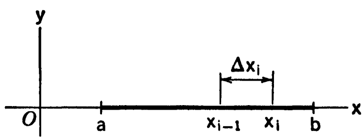



**Fig. 6-17**

We divide the wire into $n$ parts by the type of subdivision used in defining an integral (see Fig. 6-17). The mass of the piece from $x_{i-1}$ to $x_i$ is

$$\Delta M_i = \int_{x_{i-1}}^{x_i} f(x)\, dx = f(t_i)\, \Delta x_i,$$

where $t_i$ is some value of $x$ such that $x_{i-1} \leq t_i \leq x_i$. If we concentrate all the mass of this piece at $t_i$, and do likewise for all the other pieces, the moment of inertia of our auxiliary system is

$$t_1^2\, \Delta M_1 + \cdots + t_n^2\, \Delta M_n,$$

which is the same as

$$t_1^2 f(t_1)\, \Delta x_1 + \cdots + t_n^2 f(t_n)\, \Delta x_n. \qquad (3)$$

When we pass to the limit as the maximum $\Delta x_i$ approaches zero, the limit of the sum (3) is the definite integral of the function $x^2 f(x)$. Hence the moment of inertia of the wire is

$$I = \int_a^b x^2 f(x)\, dx = \int_a^b x^2 \sigma\, dx. \qquad (4)$$

Moments of inertia are often written in a form such that the mass is in evidence. The quantity

$$k = \left(\frac{I}{M}\right)^{1/2}$$

is called the *radius of gyration* for the given axis. We then have

$$I = Mk^2. \tag{5}$$

The moment of inertia of the body is the same as that of a single particle of mass $M$ concentrated at a point $k$ units from the axis.

**Example 2:** Suppose a wire $b$ feet long extends from $x = 0$ to $x = b$, with density $\sigma = x$ (mass in pounds). Find the total mass of the wire, and its moment of inertia (1) about the $y$-axis, and (2) about the line $x = b/2$.

The mass is

$$M = \int_0^b x\,dx = \frac{b^2}{2}.$$

In case (1) the moment of inertia is

$$I = \int_0^b x^3\,dx = \frac{b^4}{4} = M\frac{b^2}{2},$$

and so the radius of gyration is $k = b/\sqrt{2}$. In case (2) the moment of inertia turns out to be

$$I = \int_0^b \left(\frac{b}{2} - x\right)^2 \sigma\,dx = \int_0^b \left(\frac{b^2}{4}x - bx^2 + x^3\right) dx,$$
$$= b^4\left(\frac{1}{8} - \frac{1}{3} + \frac{1}{4}\right) = \frac{b^4}{24} = M\frac{b^2}{12}.$$

In this case, then $k = b/\sqrt{12}$.

If $I$ is the moment of inertia of a rigid body rotating about an axis with angular velocity $\omega$, the kinetic energy of the body is $\frac{1}{2}I\omega^2$. It is clear that if two bodies have the same mass and are rotating about the same axis, with the same angular velocity, that body will have the greater moment of inertia, and hence the greater kinetic energy, which has more of its mass distributed far out from the axis.

## EXERCISES

1. Consider a uniform thin rod of length $a$ extending from $x = 0$ to $x = a$. Find its moment of inertia in the form $Mk^2$ about (a) the $y$-axis; (b) the line $x = -a$; (c) the line $y = x$; (d) the line $y = x \tan\theta$, where $0 < \theta < \pi/2$. Show that in each case except (b) the moment of inertia is the same as though all the mass of the wire were concentrated at the point $(a/\sqrt{3}, 0)$.

2. (a) For the wire of Fig. 6-17 let $r$ as a function of $x$ be the perpendicular distance from the point $(x, 0)$ to a specified straight line. Write out the steps and explanations which form the basis for the formula $I = \int_a^b r^2\sigma\,dx$

which gives the moment of inertia of the wire about the given line as axis. Write $r = g(x)$, $\sigma = f(x)$.
(b) If a uniform wire of length $b$ is on the $x$-axis with its mid-point at the origin, compare its moment of inertia about a line perpendicular to the $xy$-plane and through $(-c, 0)$ (where $c > 0$) with the moment of inertia about a line perpendicular to the $xy$-plane and through $(0, c)$.

**3.** For the wire of illustrative Example 2 find the moment of inertia $I$ about the line $x = c$. Graph $I$ as a function of $c$ and find the value of $c$ for which $I$ is smallest. What is this smallest $I$?

**4.** (a) Consider a uniform triangular plate with vertices at $(0, 0)$, $(a, 0)$, $(a, b)$ in the $xy$-plane ($a$ and $b$ positive). Let it be required to formulate its moment of inertia about the $y$-axis as an integral. To do this, start by dividing the plate into thin strips parallel to the $y$-axis. Imagine each strip to be replaced by a mass particle of mass equal to the mass of the strip. By writing down the moment of inertia of the system of all these particles and passing to the limit as the maximum strip width approaches zero, you will be led to an integral giving the moment of inertia of the plate. Calculate the integral.
(b) Carry out a similar formulation and calculation for the moment of inertia of the triangle about the line $x = a$.

**5.** (a) Generalize the considerations of Exercise 4(a) so as to express as an integral the moment of inertia about the line $x = c$ of a uniform distribution of mass over the area between the curve $y = f(x)$ and the $x$-axis, $a \leq x \leq b$; assume $f(x) > 0$ if $a < x < b$.
(b) There is a uniform distribution of mass over the first quadrant area bounded by $H^2x = By^2$, $x = B$, and $y = 0$. Find the moments of inertia of this mass about the line $x = 0$, the line $x = B$, and the line $y = 0$, respectively.

**6.** Consider a homogeneous circular disk of radius $b$, thought of as occupying the area within the circle $x^2 + y^2 = b^2$. Let it be required to formulate its moment of inertia about an axis perpendicular to the $xy$-plane at the origin. Proceed as follows: First observe that if a total mass $m$ is distributed along a wire bent into the form of the circle $x^2 + y^2 = r^2$, its moment of inertia about the axis in question is $mr^2$. Then divide the disk up into a large number of thin circular rings by a series of concentric circles. If $\Delta r_i$ is the difference between the outer radius $r_i$ and inner radius $r_{i-1}$ of the $i$th ring, show that the mass of the ring is $2\pi\sigma t_i \Delta r_i$, where $t_i = \frac{1}{2}(r_i + r_{i-1})$. Then think of each ring as if it were a circular wire, and so arrive at the formulation of the moment of inertia of the disk as an integral with respect to $r$. Calculate the integral and find the radius of gyration of the disk.

## Review Questions and Problems for Chapters V and VI

### CONCEPTS AND DEFINITIONS

1. If $y = f(x)$ and $x$ is an independent variable, explain the customary meaning of the symbols $dx$ and $dy$, including the relation between them.
2. If the differential of $f$ at $x_0$ is regarded as a linear function, what graphical significance has this function, and what relation does it bear to the graph of $y = f(x)$?
3. State in words what is meant by saying that $g$ is an antiderivative of $f$ on a given interval.
4. Give the definition of "the definite integral of $f$ from $a$ to $b$," assuming that $f$ is a continuous function of $x$, defined when $a \leq x \leq b$. Begin by explaining the notation relating to approximating sums. Then explain in what sense the integral is the limit of such sums.
5. Define "upper sum" and "lower sum" in relation to the integral considered in question 4.
6. Define the arithmetic mean of the values of $f(x)$ on $[a, b]$. Illustrate by finding the arithmetic mean of $f(x) = x^3$ on $[-1, 2]$.
7. Define the concept of work, using a definite integral.
8. Define the moment of inertia of $n$ mass particles with respect to a specified axis.

### THEORY

1. Explain the circumstances in which $dy = f'(x)\,dx$ is true, even when $x$ is not an independent variable. Prove the correctness of what you say, using the chain rule.
2. Explain carefully the way in which it is shown how to express the volume of a solid of revolution by means of a definite integral.
3. State the intermediate-value theorem for continuous functions, and illustrate it graphically.
4. State the mean-value theorem pertaining to definite integrals. Learn the ideas of the proof of this theorem so thoroughly that you can write them out in your own words without having to refer to the text.
5. State and prove a theorem about the derivative with respect to $x$ of $I_a^x(f)$.
6. State and prove a theorem about finding the value of $I_a^b(f)$, using an antiderivative.
7. State and prove a theorem showing a relation between work and kinetic energy.

PROBLEMS

**1.** Identify each of the following curves as part of a parabola, and show *which* part. Calculate $y'$ and $y''$ from the parametric equations.
(a) $x = \cos 2t$, $y = \sin t$.
(b) $x = \cos t$, $y = \cos 2t$.
(c) $x = a \sin^2 t$, $y = b \cos t$.

**2.** (a) Discuss the locus defined by $x = a \cos^4 t$, $y = a \sin^4 t$. (b) Show that the line tangent to this locus at any point on it cuts the coordinate axes at points the sum of whose coordinates is $a$. (c) Show that the area between the locus and the axes is $a^2/6$.

**3.** Let $A$ be the area bounded by the line $y = mx$ and the parabola $y^2 = 2px$ (where $m$ and $p$ are $> 0$). Find the rate of change of $A$ with respect to $m$.

**4.** Let $A$ be the area bounded by the parabolas $x^2 = 2py$, $y^2 = 2qx$. Find the rate of change of $A$ when $p = 6$, $q = 2$, $p$ is decreasing 3 units per second, and $q$ is increasing 2 units per second.

**5.** Let $A$ be the area inside the triangle with vertices $(\pm 4p, 0)$, $(0, 4p)$ and above the parabola $2py = x^2$ (where $p > 0$). Find the rate of change of $A$ with respect to $p$ when $p = \frac{3}{4}$.

**6.** The area in problem 5 is revolved about the $y$-axis, thus generating a volume $V$. Calculate $V$.

**7.** A rectangle inscribed in the circle $x^2 + y^2 = a^2$ and with each side parallel to a coordinate axis has one corner at a point $(x, y)$ in the first quadrant. Express the area $A$ of the rectangle as a function of $x$, and find the arithmetic mean of the values of $A$ on the interval $[0, a]$. How does this mean value compare with the maximum value of $A$?

**8.** The area described in problem 7 is revolved about the $y$-axis, thus generating a right circular cylinder of volume $V$. Express $V$ as a function of $y$ and find the arithmetic mean of the values of $V$ on the interval $0 \leq y \leq a$. How does this mean value compare with the maximum value of $V$?

**9.** (a) Regard the ordinate of a point on the upper half of the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ as a function of $x$, and find the arithmetic mean of these ordinates for $-a \leq x \leq a$. (b) A parabola with vertex at $(0, H)$ cuts the $x$-axis at $(\pm B, 0)$ (where $B$ and $H$ are $> 0$). Find the arithmetic mean of the ordinates of this parabola (as values of a function of $x$) on the interval $|x| \leq B$.

**10.** When a stone falls freely a distance $D$ in time $T$, starting from rest, let the velocity be $v$ after falling $t$ seconds and $x$ feet, with $v = V$ when $t = T$.
(a) Regarding $v$ as a function of $t$, find its arithmetic mean for $0 \leq t \leq T$.
(b) Regarding $v$ as a function of $x$, find its arithmetic mean for $0 \leq x \leq D$.

**11.** (a) A gasoline tank has the shape of a right circular cylinder of radius $a$,

$b$ feet long (axis horizontal). When the gasoline is $h$ feet deep in the tank, find the volume of the gasoline. (b) Find the formula for the volume of gasoline in a spherical tank of radius $a$, when the gasoline is $h$ feet deep.

**12.** A cylindrical tank with vertical axis has a circular bottom of radius $b$ feet, and it is partially filled with water. A solid object is lowered slowly into the water at the rate of $c$ feet per second, with a fixed line in the solid kept vertical. Let $h$ feet be the depth of the water at the side of the tank when the lowest point of the solid is $x$ feet below the surface of the water. Let the cross-sectional area of the solid at the water surface be $A(x)$ at this time. (a) Show that the rate of change of $h$ is $cA(x)[\pi b^2 - A(x)]^{-1}$ foot per second. (b) Show that $h$ increases most rapidly when $A(x)$ is greatest. (c) What is the most rapid rate of increase of $h$ if the solid is a sphere of radius $2b/3$?

# CHAPTER VII

# FURTHER TOPICS IN ANALYTIC GEOMETRY

## 7-1 An Important Inequality

Let $a$, $b$, $u$, $v$ be any four numbers. Then

$$|au + bv| \leq (a^2 + b^2)^{1/2}(u^2 + v^2)^{1/2}. \tag{1}$$

This can be proved as follows: It is evident that $0 \leq (av - bu)^2$, and hence, as we see by expanding the square of the binomial and transposing the middle term,

$$2abuv \leq a^2v^2 + b^2u^2.$$

To both sides of this inequality let us add $a^2u^2 + b^2v^2$. We can then factor each side, and so obtain

$$(au + bv)^2 \leq (a^2 + b^2)(u^2 + v^2).$$

On taking the square root of each member of this inequality, we obtain (1).

The inequality (1) is a special case of a more general result. See Exercise 1 at the end of this section.

As a consequence of (1) we can prove the following assertion:

*If $a$ and $b$ are fixed and not both zero, and if $u$ and $v$ vary arbitrarily except for the requirement that $au + bv = c$, where $c$ is fixed, then the minimum possible value of $(u^2 + v^2)^{1/2}$ is $|c|(a^2 + b^2)^{-1/2}$.*

To prove this, we start from $au + bv = c$ and apply (1), obtaining

$$|c| \leq (a^2 + b^2)^{1/2}(u^2 + v^2)^{1/2},$$

or

$$(u^2 + v^2)^{1/2} \geq \frac{|c|}{(a^2 + b^2)^{1/2}}. \tag{2}$$

Hence it is certain that $(u^2 + v^2)^{1/2}$ can never be *less* than $|c|(a^2 + b^2)^{-1/2}$. But, if we choose

$$u = \frac{ac}{a^2 + b^2}, \qquad v = \frac{bc}{a^2 + b^2},$$

we find that

$$au + bv = \frac{a^2c}{a^2 + b^2} + \frac{b^2c}{a^2 + b^2} = c$$

and

$$u^2 + v^2 = \frac{a^2c^2}{(a^2 + b^2)^2} + \frac{b^2c^2}{(a^2 + b^2)^2} = \frac{c^2}{a^2 + b^2},$$

so that in this case (2) becomes an equality. Hence the value $|c|(a^2 + b^2)^{-1/2}$ is actually attained as a minimum. This proves the assertion.

This result will be used in a geometric way in the next section.

### EXERCISES

1. Prove that $|au + bv + cw| \leq (a^2 + b^2 + c^2)^{1/2}(u^2 + v^2 + w^2)^{1/2}$, using the following suggestion. Let $P(t) = (a + tu)^2 + (b + tv)^2 + (c + tw)^2$ and expand this into a form $At^2 + Bt + C$. Explain why it is necessary that $B^2 - 4AC \leq 0$, and show that this leads to the desired proof. Then state a generalization involving a sum of four or more products, and prove it.
2. If $u$, $v$, and $w$ vary arbitrarily except for the requirement that $u^2 + v^2 + w^2 = 1$, prove that the largest possible value of $|au + bv + cw|$ is $(a^2 + b^2 + c^2)^{1/2}$.

## 7-2 The Distance Between a Point and a Line

As we know (see § 1-3, and especially Theorem 1-A), an equation of the form $Ax + By + C = 0$ in which $A$ and $B$ are not both zero has a straight line as its graph. It is of interest to know how to express the perpendicular distance from any given point $(x_0, y_0)$ to the line in terms of $x_0$, $y_0$ and the coefficients $A$, $B$, $C$. Let this distance be $d$. We shall show that

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}. \tag{1}$$

We prove this by using the result stated in § 7-1. By transposing $C$ and subtracting $Ax_0 + By_0$ from both sides, we can write the equation of the line in the form

$$A(x - x_0) + B(y - y_0) = -(Ax_0 + By_0 + C). \tag{2}$$

Let $u = x - x_0$, $v = y - y_0$, $a = A$, $b = B$, $c = -(Ax_0 + By_0 + C)$. Then (2) takes the appearance $au + bv = c$. Now $d$ is the smallest possible value of

$$[(x - x_0)^2 + (y - y_0)^2]^{1/2} = (u^2 + v^2)^{1/2}$$

as $(x, y)$ varies in such a way that (2) is satisfied (see Fig. 7-1). Hence we can apply the result stated in italics in § 7-1, and we get the expression for $d$ in (1) as the desired minimum value.

**Example 1:** Find the altitude of the triangle with vertices (2, 4), (−2, −4), (8, 1), considering the two last mentioned vertices as ends of the base.

The equation of the base line is $y + 4 = \frac{1}{2}(x + 2)$, or $x - 2y - 6 = 0$. The perpendicular distance between the vertex (2, 4) and this line is

$$d = \frac{|1(2) - 2(4) - 6|}{\sqrt{1 + 4}} = \frac{12}{\sqrt{5}}.$$

This is the required altitude.

If the equations of two intersecting lines are given, we can find equations of the lines which bisect the angles between the lines by using the fact

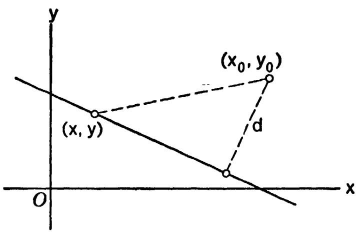



**Fig. 7-1**

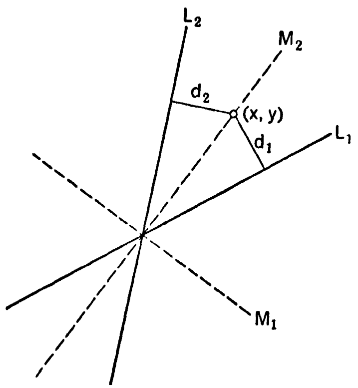



**Fig. 7-2**

that if a point moves so as to remain at equal distances from the two lines, it must remain on one of the angle bisectors (see Fig. 7-2). Suppose the lines are $L_1$, $L_2$, with equations

$$A_1x + B_1y + C_1 = 0, \qquad A_2x + B_2y + C_2 = 0.$$

Now let the point $(x, y)$ move on one of the bisectors. Its distance from $L_1$ is

$$d_1 = \frac{|A_1x + B_1y + C_1|}{\sqrt{A_1^2 + B_1^2}},$$

and there is a similar expression for $d_2$. Hence the equation

$$\frac{|A_1x + B_1y + C_1|}{\sqrt{A_1^2 + B_1^2}} = \frac{|A_2x + B_2y + C_2|}{\sqrt{A_2^2 + B_2^2}}$$

expresses the fact that $(x, y)$ is on one or the other of the two bisectors. There are now two possibilities, which we express with the ± sign:

$$\frac{A_1x + B_1y + C}{\sqrt{A_1^2 + B_1^2}} = \pm\frac{A_2x + B_2y + C}{\sqrt{A_2^2 + B_2^2}}. \tag{3}$$

By choosing first the + and then the − sign, we obtain the equations of two lines; these are the two bisectors $M_1$ and $M_2$. We can distinguish one bisector from the other in particular problems by graphing the two original lines and then examining the slopes of all four lines.

**Example 2:** Find equations of the bisectors of the angles formed by the lines $L_1$, $L_2$ whose equations are $3x - 4y + 8 = 0$, $5x + 12y - 26 = 0$.
According to the method, we have

$$\frac{3x - 4y + 8}{5} = \pm\frac{5x + 12y - 26}{13}.$$

On reducing, we find the equations of the bisectors to be

$$7x - 56y + 117 = 0 \tag{4}$$

and

$$32x + 4y - 13 = 0. \tag{5}$$

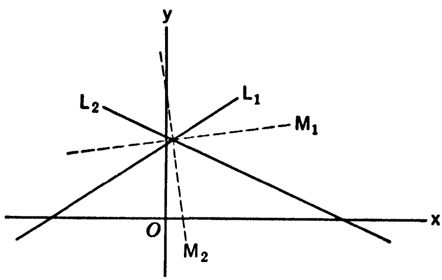



**Fig. 7-3**

Figure 7-3 shows the situation. The slopes of $L_1$ and $L_2$ are, respectively, $\frac{3}{4}$ and $-\frac{5}{12}$. The slopes of the bisectors are $\frac{7}{56}$ [from equation (4)] and $-8$ [from equation (5)]. Clearly then from the figure, (4) is the equation of $M_1$ and (5) is the equation of $M_2$.

## EXERCISES

1. Find the perpendicular distances from the point (10, 14) to the legs of the isosceles triangle with vertices at (0, 16), (±10, 0).

2. Find the perpendicular distances between the following pairs of parallel lines.
   (a) $4x + 3y = 9$ and $8x + 6y = 11$.
   (b) $x - 2y + 10 = 0$ and $2x - 4y = 3$.
   (c) $x - 3y = 4$ and $6y = 2x + 15$.

**3.** The two lines $3x + 4y = 9$, $12x - 5y = 12$ divide the plane into four parts, one of which contains the origin. (a) Find the equation of that bisector of the angles between the lines which penetrates the aforementioned part of the plane. (b) Find the points in which the other bisector cuts the coordinate axes.

**4.** (a) Find the equations of the bisectors of the interior angles of the triangle with vertices at $(1, -8)$, $(9, -2)$, $(1, 13)$. (b) Find the point of intersection of these bisectors. (c) Find the radius of the circle inscribed in the triangle.

**5.** Proceed as directed in Exercise 4 for the case of the triangle with vertices at $(-8, -4)$, $(13, -4)$, $(8, 8)$.

**6.** Find two points on the $x$-axis, each a distance 4 from the line $12x + 5y = 20$.

**7.** Find two points on the line $x + y = 1$, each a distance $\sqrt{13}$ from the line $3y = 2x$.

**8.** Find two points which are equidistant from the points $(4, 8)$, $(9, 3)$ and 5 units from the line $4x - 3y + 13 = 0$.

**9.** A point moves in such a way that its distance from $(1, 1)$ is the same as its distance from the line $x + y + 2 = 0$. Make a diagram and sketch the locus. What is it called? (a) Find the equation of the locus in a form free of radicals. (b) Find the point of the curve at which the tangent is parallel to the $x$-axis. (c) At what point is the tangent parallel to the $y$-axis?

**10.** Suppose the line $L$ does not go through the origin. Let $N$ be the point of intersection of $L$ and the line through 0 perpendicular to $L$. Let $\alpha$ be the angle from the positive $x$-axis to the ray $ON$, and let $p$ be the length $ON$ (see Fig. 7-4).

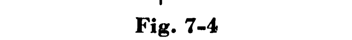


**Fig. 7-4**

(a) Show that the equation of $L$ is $x \cos \alpha + y \sin \alpha - p = 0$. This is called the *normal form* of the equation of $L$.

(b) If $Ax + By + C = 0$ is another form of the equation of $L$, show that we can convert this to the normal form by dividing by $\sqrt{A^2 + B^2}$ if $C < 0$, and by dividing by $-\sqrt{A^2 + B^2}$ if $C > 0$.

(c) Convert each of the following equations to normal form, and find $p$.

(i) $4x - 3y + 7 = 0$. (iii) $15x + 8y - 51 = 0$.

(ii) $12y - 5x = 26$. (iv) $x = 3y - 5$.

(d) If the equation of line $L$ is put in normal form, and $(x_1, y_1)$ is a point not on $L$, show that $x_1 \cos \alpha + y_1 \sin \alpha - p$ is positive if $(x_1, y_1)$ is on the side of $L$ opposite to that of the origin, and negative if $(x_1, y_1)$ and the origin are on the same side of $L$. Show also that the distance from the point to the line is $d = |x_1 \cos \alpha + y_1 \sin \alpha - p|$.

(e) If two lines are given, in normal forms $x \cos \alpha_1 + y \sin \alpha_1 - p_1 = 0$, $x \cos \alpha_2 + y \sin \alpha_2 - p_2 = 0$, what is the equation of that bisector of the angles between these lines which goes into the angle where the origin lies? What is the equation of the other angle bisector?

**11.** Draw a figure showing the three lines $12x - 5y + 15 = 0$, $3x - 4y = 3$, $3x + 4y = 3$. How many circles are there which are tangent to all three lines? Construct them, roughly. Then use the methods of this section to find the center and radius of each circle. See Exercise 10, especially parts (d) and (e).

**12.** Derive the formula (1) for the distance $d$ from the point $(x_0, y_0)$ to the line $L$ by using calculus to find the minimum value of $(x - x_0)^2 + (y - y_0)^2$ when the point $(x, y)$ varies along $L$. This is an extremal problem with the equation of $L$ expressing the side condition on $x$ and $y$.

## 7-3 Families of Lines

Sometimes it is useful to consider all the straight lines which have some geometric property in common. For example, we might consider all lines through (1, 2), or all lines with slope $-\frac{3}{4}$, or all lines which form with the coordinate axes a triangle of area 24. Any well-determined collection of lines is called a *family* of lines. Some ways in which families of lines are of interest will appear in the following illustrative examples.

**Example 1:** Describe by equations the family of all lines which form with the coordinate axes a triangle of area 24. Find all such lines which pass through the point (6, 4).

We use the equation of a line in the intercept form

$$\frac{x}{a} + \frac{y}{b} = 1;$$

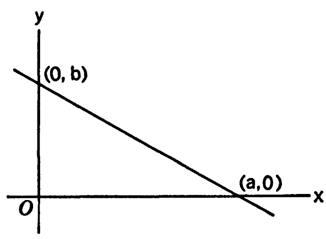



**Fig. 7-5**

see § 1-4 and Fig. 7-5. There are two cases to consider, according as $a$ and $b$ are of the same or of opposite signs. In the first case, the area of the triangle is $ab/2$, and so we must have $ab = 48$, or $b = 48/a$. Hence we can write the equation of this part of the family of lines in the form

$$\frac{x}{a} + \frac{ay}{48} = 1.$$

None of these lines go through (6, 4). For, if we put $x = 6$, $y = 4$, and try to solve for $a$, we get a quadratic with imaginary roots:

$$\frac{6}{a} + \frac{4a}{48} = 1, \quad a^2 - 12a + 72 = 0, \quad a = 6(1 \pm i).$$

In the second case the area of the triangle is $-ab/2$, so $ab = -48$, or $b = -48/a$. In this case the equation of the part of the family of lines is

$$\frac{x}{a} - \frac{ay}{48} = 1.$$

One of these lines will pass through (6, 4) if $a$ is determined as follows:

$$\frac{6}{a} - \frac{4a}{48} = 1, \quad a^2 + 12a - 72 = 0, \quad a = -6(1 \pm \sqrt{3}).$$

Thus there are two lines which meet the requirements laid down at the outset.

In many simple but interesting situations a family of lines is defined by an equation in $x$ and $y$ which involves one or more auxiliary variables, or parameters, like $a$ in the foregoing example. If some geometric condition is imposed which serves to select out a particular member of the family, the application of this condition may be expressible as an equation of condition from which the parameter values may be determined, and thus we may find an equation of the particular line that is wanted.

**Example 2:** Find an equation which describes the family of all lines through the intersection of the two lines

$$2x + 5y = 10, \qquad 3x - 4y = 12. \tag{1}$$

Then find the particular member of the family which passes through (2, 2).

Consider the equation

$$h(2x + 5y - 10) + k(3x - 4y - 12) = 0, \tag{2}$$

where $h$ and $k$ are arbitrary parameters, not both zero. When $h$ and $k$ are fixed, this is a linear equation, and hence it represents a line. Moreover, this line goes through the intersection of the two given lines. For if the point $(x, y)$ satisfies both equations (1), it certainly satisfies equation (2). By varying $h$ and $k$ we obtain many different lines. We can rewrite (2) in the form

$$(2h + 3k)x + (5h - 4k)y - 10h - 12k = 0. \tag{3}$$

In this form we see that we can choose $h$ and $k$ so as to obtain any specified line of the family under consideration. For if $A$ and $B$ are any numbers, not both zero, we can solve for $h$ and $k$ in the equations

$$2h + 3k = A, \qquad 5h - 4k = B,$$

and thus throw (3) into the form

$$Ax + By + C = 0,$$

with $A$ and $B$ prescribed ahead of time. This means that every possible slope is obtainable, and so we get all members of the family in this way. In order to get the particular line which goes through (2, 2), we put $x = y = 2$ in (2). Then

$$4h - 14k = 0, \quad \text{or} \quad 2h - 7k = 0. \tag{4}$$

We do not expect to find unique values for $h$ and $k$; it is only their ratio which is determined. We can choose any nonzero value of $k$ in (4) and then solve for $h$. We take $k = 2$ and get $h = 7$; then (3) becomes

$$20x + 27y - 94 = 0.$$

**Example 3:** The family of all tangents to a given curve is often quite interesting. We shall exhibit the family of all tangents to the parabola $y = x^2$. Let $(\alpha, \beta)$ be the point of tangency, so that $\beta = \alpha^2$. We shall use $\alpha$ as the parameter in describing the family. From $dy/dx = 2x$ we see that the slope at $(\alpha, \beta)$ is $2\alpha$. Hence the equation of the tangent is

$$y - \alpha^2 = 2\alpha(x - \alpha), \quad \text{or} \quad y = 2\alpha x - \alpha^2.$$

This equation shows, for instance, that the $y$-intercept of each tangent is the negative of the $y$-coordinate of the point of tangency. If we wish to find the

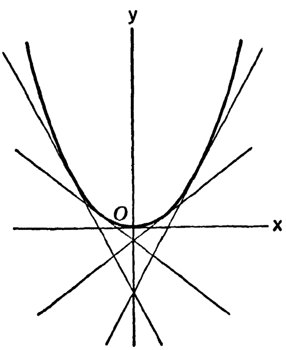



**Fig. 7-6**

tangent which goes through a particular point $(x_0, y_0)$ in the plane, we try to find $\alpha$ so that $y_0 = 2\alpha x_0 - \alpha^2$. This quadratic has solutions

$$\alpha = \frac{2x_0 \pm \sqrt{4x_0^2 - 4y_0}}{2} = x_0 \pm \sqrt{x_0^2 - y_0}.$$

There are two solutions if $x_0^2 - y_0 > 0$; there is one solution if $x_0^2 - y_0 = 0$; and there is no solution (since $\alpha$ must be real) if $x_0^2 - y_0 < 0$. Observe that $y_0 > x_0^2$ means that the point $(x_0, y_0)$ is inside the parabola. These findings agree with what we expect from looking at Fig. 7-6.

## EXERCISES

**1.** By using a parameter in an appropriate way, write an equation to describe each of the following families of lines. Then pick out the particular line of the family that satisfies the geometric condition as stated. Draw a figure showing several lines of the family.

(a) All lines with $y$-intercept 4. The particular one which is perpendicular to the line $3x + 2y + 6 = 0$.

(b) All lines parallel to the line $2y = x + 3$. The particular one through the point $(-2, 1)$.

(c) All lines perpendicular to the line $2x + 5y = 10$. The particular ones which make with the coordinate axes a triangle of area 4.

(d) All lines at a distance of 5 units from the origin. The particular one with positive slope and $y$-intercept $\frac{25}{4}$. See Exercise 10 in § 7-2.

(e) All lines with $x$-intercept $-3$. The particular one which is parallel to the line $3x + y + 6 = 0$.

**2.** What geometric feature is shared by all the lines in each of the following families? Find the line or lines of the family which has the property stated after the equation of the family.

(a) $3x - 2y + C = 0$. Through $(1, 4)$.

(b) $4x - 2ky + 8 = 0$. Parallel to $x = y$.

(c) $kx + \sqrt{1 - k^2}y - 6 = 0$. Slope 2.

(d) $\dfrac{x}{a} + \dfrac{y}{\sqrt{25 - a^2}} = 1$. Intercept 4 on the $y$-axis.

(e) $y = m(x - 3) + 5$. Perpendicular to $2y - 3x = 4$.

(f) $y - mx = 2m$. Distance from the origin to the line is 1 unit.

**3.** In each case write an equation for the family of lines through the intersection of the given pair of lines. Then find that particular line or lines which fulfill the added condition.

(a) $3x + 2y = 58$, $2x = 5y - 50$. Parallel to $3x - 2y + 8 = 0$.

(b) $x + y = 9$, $3x - 4y + 1 = 0$. Perpendicular to $2x + 3y = 10$.

(c) $x - 2y = 3$, $3x - y = 2$. Through the intersection of $x + y = 2$ and $x - y = 6$. Can you solve the problem without finding the intersection of either pair of lines?

(d) $2x - 11y = 15$, $7x - y = 30$. Tangent to the circle $x^2 + y^2 = 9$.

**4.** (a) Write an equation for the family of all lines with 8 as the algebraic sum of the $x$- and $y$-intercepts. (b) Pick out the ones with slopes $-1$, $\frac{1}{2}$, 2, and draw them. (c) Are all slopes possible for the lines of the family? (d) Show how to determine the parameter so that the line goes through $(x_0, y_0)$, if such is possible. What equation must $(x_0, y_0)$ satisfy if there is exactly one line of the family through it?

**5.** If $(x, y)$ moves so that its distance from the line of slope $-1$ through the origin is equal to its distance from the point $(4, 4)$, show that the equation of its locus is $x^2 + y^2 - 2xy - 16x - 16y + 64 = 0$. This is a parabola, of course. Draw it. Show that, if $L$ is a line tangent to this curve at any point, the algebraic sum of the intercepts of $L$ on the axes is 8. See Exercise 4.

**6.** (a) Show that, in the first case of Example 1 (i.e., where $ab = 48$) there are two lines of this part of the family through the point (3, 2), and one line through (3, 4).
(b) Show that there are two of these lines through the point $(x_0, y_0)$ if and only if the point lies in the part of the plane between the two branches of the rectangular hyperbola $xy = 12$.

**7.** Write an equation of the family of all tangents to the hyperbola $xy = b^2$.

**8.** (a) Write an equation of the family of all tangents to the parabola $y^2 = 2px$. (b) Show that, if each tangent is paired with the one which is perpendicular to it, the two points of tangency and the focus of the parabola are collinear.

**9.** (a) Find the family of all tangents to the first quadrant portion of the curve $x^{2/3} + y^{2/3} = b^{2/3}$. (b) Show that the length of each tangent cut off by the axes is $b$.

**10.** The curves $C_1 : x^2 = 2py$ and $C_2 : 27px^2 = 8(y - p)^3$ stand in an interesting relation to each other, namely, the family of all normals to $C_1$ is the same as the family of all tangents to $C_2$. Prove this as follows: If $(\alpha, \beta)$ is a point of $C_1$, write the equation of the normal to $C_1$ at this point, keeping $\alpha$ as the parameter. Then, by calculating the slope of $C_2$ at an arbitrary point, find the point of $C_2$ at which its tangent coincides with the aforementioned normal to $C_1$. Then observe that, as $(\alpha, \beta)$ traces out $C_1$, this other point traces out $C_2$. It will be helpful if you begin by making a fairly good graph of $C_1$ and $C_2$ on the same axes.

## 7-4 Families of Circles

The family of all circles which pass through two given points has many features of interest. Figure 7-7 shows some of the circles of this family for

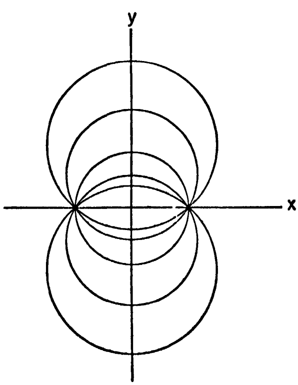



Fig. 7-7

the case in which the two points are $(b, 0)$ and $(-b, 0)$. Any circle of this family must have its center at some point $(0, c)$ on the $y$-axis. Its radius must then be $(b^2 + c^2)^{1/2}$, so its equation is

$$x^2 + (y - c)^2 = b^2 + c^2, \quad \text{or} \quad x^2 + y^2 - 2cy - b^2 = 0. \tag{1}$$

This is a one parameter family of circles, the parameter being $c$. We consider $b$ as fixed.

There is another family of circles which is related in a most interesting way to the family just described. Some of the circles of this second family enclose the point $(b, 0)$, and the others enclose the point $(-b, 0)$. The most

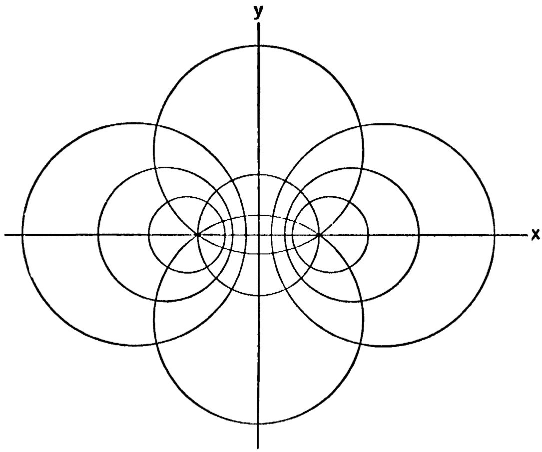



**Fig. 7-8**

remarkable feature of these two families of circles, when we consider both families at once, is that each circle of one family is *orthogonal* to every circle of the other family; that is, where a circle of one family intersects a circle of the other family, the tangents to the two circles are perpendicular (see Fig. 7-8). This second family is defined by the equation

$$(x - a)^2 + y^2 = a^2 - b^2, \quad \text{or} \quad x^2 - 2ax + y^2 + b^2 = 0. \tag{2}$$

Here $a$ is the parameter, and the center of the circle, for a particular $a$, is at $(a, 0)$. Note that we require $a^2 > b^2$.

In order to verify that the two families of circles are orthogonal in the sense described, we observe the following: Two circles, of radii $r_1$ and $r_2$, with distance $h$ between their centers, are orthogonal if and only if

$$r_1^2 + r_2^2 = h^2. \tag{3}$$

For the proof of this, see Fig. 7-9.

Now, to prove that each of the circles (1) is orthogonal to each of the circles (2), observe that the respective radii are $(b^2 + c^2)^{1/2}$ and $(a^2 - b^2)^{1/2}$, and that the distance between centers is $(a^2 + c^2)^{1/2}$. We see at once that condition (3) is fulfilled, so we do indeed have orthogonality.

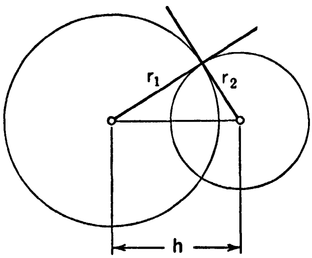



Fig. 7-9

There is a physical situation which gives rise to this pair of families of circles. Imagine two parallel wires piercing the $xy$-plane at right angles, one at $(b, 0)$ and one at $(-b, 0)$. Suppose the wires carry static electric charges distributed uniformly along their lengths, the amount per unit lengths for the two wires being equal in size but opposite in sign. For definiteness, suppose the wire through $(-b, 0)$ carries the positive charge. Then a field of electrostatic force is produced. At any point in space the direction of the force is in a plane perpendicular to the wires. Since the situation is the same in all such planes, we may as well examine the $xy$-plane. It turns out to be true (though we cannot explain the physical reasons here) that the direction of the force is always from the positively charged wire to the negatively charged wire *along the circle of the family* (1) *through the particular point.* The circles of the other family arise from a consideration of the electrostatic *potential.* This is a term whose meaning we shall not try to define here. With each point in the electrostatic field there is associated a number, called the potential at that point. One main significance of the potential is that it enables one to calculate the work required to move a unit positive charge from one point to another in the field. Now, around each wire there is a family of cylindrical surfaces, on each of which the potential is constant. These cylinders intersect the $xy$-plane in the circles of the family (2). The electrostatic force at a point on one of these cylinders is perpendicular to the cylinder. This is the physical realization of the orthogonality of the two families of circles.

### *The Radical Axis of Two Circles*

Consider a given circle with center $(a, b)$ and radius $r$, and let $(x_1, y_1)$ be a point outside this circle. Draw a line from $(x_1, y_1)$ tangent to the circle,

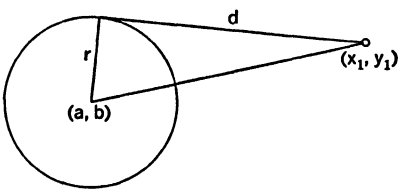



Fig. 7-10

and let $d$ be the distance from $(x_1, y_1)$ to the point of tangency (see Fig. 7-10). We call $d$ the length of the tangent from the point to the circle. It is clear from the figure that

$$d^2 = (x_1 - a)^2 + (y_1 - b)^2 - r^2. \tag{4}$$

**Example 1:** The length of the tangent from (8, 5) to the circle of radius 3 with center at (1, 2) is

$$d = [(8 - 1)^2 + (5 - 2)^2 - 9]^{1/2} = 7.$$

Now consider any two given circles $C_1$, $C_2$ with different centers. The circles may or may not intersect. There is a certain straight line $L$ with the property that $L$ is the locus of all points $P$ for which the length of the tangent from $P$ to $C_1$ is the same as the length of the tangent from $P$ to $C_2$. This line is called the *radical axis* of the two circles.

In order to see that there is such a line, let us suppose $C_1$ has center $(a_1, b_1)$ and radius $r_1$, with similar notations pertaining to $C_2$. Then, if $P$ has coordinates $(x, y)$, we see from (4) that the condition which $P$ must satisfy is

$$(x - a_1)^2 + (y - b_1)^2 - r_1^2 = (x - a_2)^2 + (y - b_2)^2 - r_2^2. \tag{5}$$

On simplifying this equation, we see that it is equivalent to the equation

$$2(a_2 - a_1)x + 2(b_2 - b_1)y + a_1^2 - a_2^2 + b_1^2 - b_2^2 + r_2^2 - r_1^2 = 0. \tag{6}$$

Since the circles have different centers, the coefficients of $x$ and $y$ in (6) are not both zero, and so we have the equation of a straight line; this line is the radical axis. Equation (6) should not be memorized, but we should notice how the equation may be obtained when we know the equations of the two circles. We merely write the equation of each circle in the form

$$x^2 + y^2 + Ax + By + C = 0,$$

and subtract one equation from the other.

**Example 2:** Find the radical axis of the two circles $(x - 2)^2 + y^2 = 9$, $(x - 8)^2 + y^2 = 25$.

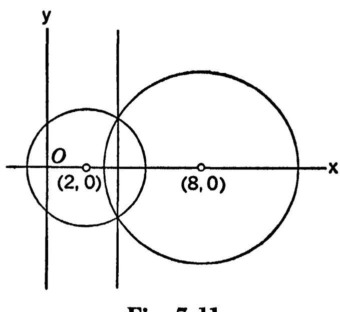



**Fig. 7-11**

We write

$$\begin{array}{r} x^2 - \phantom{1}4x + \phantom{6}4 + y^2 - \phantom{2}9 = 0 \\ x^2 - 16x + 64 + y^2 - 25 = 0 \\ \hline 12x - 60 \qquad + 16 = 0 \end{array}, \quad \text{or} \quad x = 11/3.$$

In this case the two circles intersect, and the radical axis goes through the points of intersection. See Fig. 7-11.

What was true in Example 2 is true in general: *If two circles intersect, their radical axis is the line through their two points of intersection.* On the other hand, if two circles do not intersect and are not concentric, their radical axis does not intersect either circle. In the case of two tangent circles, the radical axis is the common tangent line at the point where the circles are tangent.

*Coaxal Families of Circles*

We introduce the general idea of this topic by considering a particular example.

**Example 3:** The equations of the two circles in Example 2 can be written in the forms

$$x^2 + y^2 - 4x - 5 = 0, \qquad x^2 + y^2 - 16x + 39 = 0. \tag{7}$$

Let us write the equation

$$(1 - k)(x^2 + y^2 - 4x - 5) + k(x^2 + y^2 - 16x + 39) = 0, \tag{8}$$

where $k$ is an arbitrary constant. This equation can be rewritten as

$$x^2 + y^2 - (4 + 12k)x - 5 + 44k = 0. \tag{9}$$

This is the equation of a circle. We see from (7) and (8) that if a point $(x, y)$ lies on both of the two original circles, it also lies on the circle defined by (9). Hence (9) is the equation of a circle which goes through the two points of intersection of the given circles. By completing the square in (9) we can see that the center of the circle is at the point $(a, 0)$, where

$$a = 2 + 6k, \quad \text{or} \quad k = \frac{a - 2}{6}. \tag{10}$$

From this we see that, when $a$ is assigned, we can choose $k$ so that (10) will hold. Hence (9) may be considered as the equation of the family of all circles which pass through the two points of intersection of the original pair of circles. Any pair of these circles has the same radical axis, namely the line $x = \frac{11}{3}$.

A family of circles is called a *coaxal family* if all pairs of circles in the family have the same radical axis. The following general proposition can be proved:

*Let $C_1$ and $C_2$ be two nonconcentric circles, and let their equations be denoted for convenience in the abbreviated form $f_1(x, y) = 0$, $f_2(x, y) = 0$, where in each case the coefficient of $x^2 + y^2$ is 1. Then the equation*

$$(1 - k)f_1(x, y) + kf_2(x, y) = 0, \tag{11}$$

*in which $k$ is a parameter, is a coaxal family consisting of all circles which, when paired with either $C_1$ or $C_2$, yield the same radical axis as the pair $C_1$, $C_2$.*

If $C_1$ and $C_2$ intersect, then $k$ may be assigned any value whatever. If $C_1$ and $C_2$ do not intersect, there are certain values of $k$ for which (11) has no locus (for reasons explained in § 3-8). In this case the coaxal family consists of two parts; the circles of one part of the family all enclose a certain point, and the circles of the other part enclose the point which is the mirror image of this first point relative to the common radical axis. Figure 7-8 illustrates the two general types of coaxal families.

**Example 4:** Let $C_1$ and $C_2$ have centers at (5, 0) and (−5, 0), respectively, and equal radii $r = 3$. Their equations are

$$x^2 + y^2 - 10x + 16 = 0, \qquad x^2 + y^2 + 10x + 16 = 0,$$

and the coaxal family is

$$(1 - k)(x^2 + y^2 - 10x + 16) + k(x^2 + y^2 + 10x + 16) = 0,$$

or

$$x^2 + y^2 + 10(2k - 1)x + 16 = 0. \tag{12}$$

For a given $k$ the center is at $[5(2k + 1), 0]$ and the radius is $[25(2k + 1)^2 - 16]^{1/2}$. The circle reduces to a point if $5(2k - 1) = \pm 4$. The two points thus determined are $(\pm 4, 0)$. In this case the common radical axis is the line $x = 0$. There is no locus for (12) if $25(2k - 1)^2 - 16 < 0$. This inequality is equivalent to $|2k - 1| < \frac{4}{5}$, which in turn is equivalent to the double inequality $\frac{1}{10} < k < \frac{9}{10}$.

## EXERCISES

**1.** (a) Find the length of the tangents from (4, 1) to the circle $x^2 + y^2 + 6x - 4y - 12 = 0$.
(b) What are the points of tangency?

**2.** Find the intersection of the two circles $x^2 + y^2 + 3x - 3y = 52$, $x^2 + y^2 - 2x + 2y = 32$.

**3.** (a) Find the radical axis of the two circles $x^2 + (y - 2)^2 = 4$, $x^2 + (y - 3)^2 = 11$.
(b) Write an equation of the largest coaxal family which includes these two circles.
(c) Find the members of this family with centers at (0, −4), (0, −5), $(0, \frac{3}{2})$, and $(0, -\frac{7}{2})$, respectively.
(d) Find the two points onto which members of the coaxal family contract as their radii shrink to zero.

**4.** Find the circle through the point of intersection of the circles $x^2 + y^2 + 8x - 4y = 8$ and $x^2 + y^2 + 6x - 4y = 14$ and having its center on the $y$-axis.

**5.** Find the points from which the tangents to each of the following three circles are of equal length: $x^2 + y^2 + 8x + 6y = 0$, $x^2 + y^2 + 9x + 3y = 1$, $x^2 + y^2 + 7x + 4y = -9$.

**6.** Find a circle coaxal with each of the circles $x^2 + y^2 - 2y - 1 = 0$, $x^2 + y^2 + 4y - 1 = 0$ and going through the point $(4, 2)$.

**7.** Find the circle through the point $(1, -1)$ and the points of intersection of the circles $x^2 + y^2 + 4x - 2y + 1 = 0$, $x^2 + y^2 + 2x - 8y + 8 = 0$.

**8.** For the coaxal family (12):
(a) Show that for each point $P$ of the plane not on the $y$-axis there is exactly one circle of the family passing through $P$;
(b) find the center and radius of the member of the family through $(5, 2)$;
(c) find the center and radius of the member of the family through $(-6\sqrt{2}, 2\sqrt{2})$.

**9.** (a) Find the radical axis of the circles $x^2 + y^2 - 6x - 4y + 5 = 0$, $x^2 + y^2 + 8x - 2y = 1$, and prove that it is tangent to both of them. Draw a figure.
(b) Find the circle coaxal with these two circles and going through the origin.

**10.** Find the family of all circles which share the line $x + 2y = 20$ as radical axis with the circle $x^2 + y^2 - 4x + 2y = 75$.

**11.** Prove analytically that the radical axis of two circles is perpendicular to their line of centers.

**12.** Prove the proposition stated in the text in connection with (11). *Suggestion:* The radical axis of $C_1$ and $C_2$ has equation $f_1(x, y) - f_2(x, y) = 0$. First show that any two distinct members of the family (11) have this same equation for their radical axis. Observe that the equation is essentially unchanged if multiplied through by a nonzero constant. Then show that, if $C$, with equation $f(x, y) = 0$, is any circle such that $C$ and $C_1$ have the same radical axis as $C_1$ and $C_2$, then $f(x, y)$ is of the form $(1 - k) f_1(x, y) + k f_2(x, y)$, where $k$ is some constant.

**13.** Consider three circles $C_1, C_2, C_3$, no two of them concentric. Let $L_1, L_2, L_3$ be the radical axes of the pairs $(C_2, C_3)$, $(C_3, C_1)$, $(C_1, C_2)$, respectively. Show that, if two of these lines intersect, the third one goes through the point of intersection of the other two.

**14.** Given a circle $C$, a line $L$, and a constant $t > 0$, consider the locus of a point $P$ outside $C$ which moves in such a way that $\overline{PT}^2 = t\overline{PN}$, where $PT$ is a tangent from $P$ to $C$ and $PN$ is a perpendicular from $P$ to $L$. Show that $P$ moves on a circle $C_1$ such that the radical axis of $C$ and $C_1$ is $L$.

**15.** Suppose $b > 0$. Let $P_1 = (b, 0)$, $P_2 = (-b, 0)$. If $0 < t \neq 1$, consider the locus of points $P$ such that $\overline{PP_1} = t\overline{PP_2}$. (a) Show that this is a circle which encloses $P_1$ if $t < 1$ and $P_2$ if $1 < t$. (b) Show that the family of all such circles, when $t$ is regarded as a parameter, is a coaxal family. (c) Two points are said to be mutually *inverse* with respect to a given circle if they lie on the same ray extending from the center of that circle, and if the product of their distances from that center is equal to the square of the radius. Show that the two points $P_1$, $P_2$ here are mutually inverse with respect to every member of the coaxal family.

## 7-5 Confocal Ellipses and Hyperbolas

Let us consider the family of all ellipses which have the points $(\pm c, 0)$ as foci. If the lengths of the major and minor axes of one of these ellipses are $2a$ and $2b$, we know that $a^2 = b^2 + c^2$ (see § 3-8). Let us write $h$ in place of $b^2$ and regard $h$ as a parameter which can have any positive value. Then $a^2 = c^2 + h$, and the equation

$$\frac{x^2}{c^2 + h} + \frac{y^2}{h} = 1 \tag{1}$$

describes the family of confocal ellipses. When $h$ is small, the ellipse is narrow, fitting closely around the line segment joining the foci. When $h$ is large, the ellipse is large and nearly circular (see Fig. 7-12). Any point not on the line joining the foci is on exactly one of these ellipses.

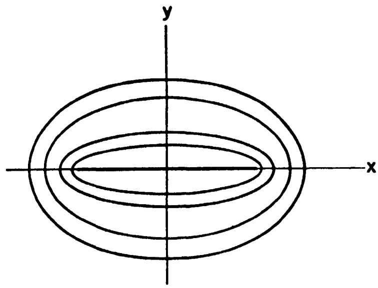


Fig. 7-12

In connection with this family of confocal ellipses it is interesting to consider the family of all hyperbolas having the same foci as the ellipses. If the vertices of one of these hyperbolas are at $(\pm a, 0)$, we know that $a^2 < c^2$. Let us write $a^2 = c^2 + k$, where $-c^2 < k < 0$. Then, as we know from § 3-9, the equation of the hyperbola is

$$\frac{x^2}{c^2 + k} + \frac{y^2}{k} = 1; \tag{2}$$

hence, with $k$ as the parameter, (2) is the equation of the family of confocal hyperbolas. For $k$ close to zero, the branches of the hyperbola fit very close around the foci and the outer extremities of the $x$-axis. As $k$ gets close to $-c^2$, the branches approach the $y$-axis. As we shall see presently, each hyperbola cuts all the confocal ellipses at right angles (see Fig. 7-13). Any point not on the $y$-axis, and not on the $x$-axis with $|x| > c$, is on exactly one of these hyperbolas.

**Example 1:** Take $c = 5$, and find $h$ and $k$ so as to get the ellipse and the hyperbola through the point (8, 4).

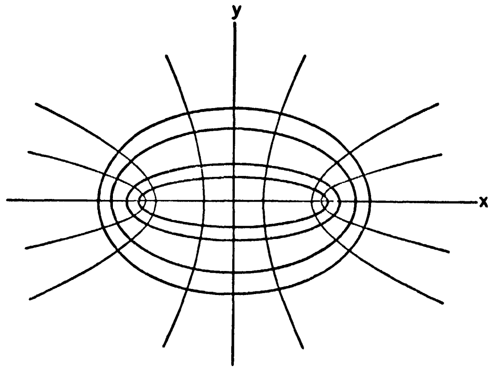



**Fig. 7-13**

We observe that equations (1) and (2) have the same appearance, except that $h > 0$, while $k < 0$. If we put $x = 8$ and $y = 4$ in (1), we get a quadratic equation for $h$:

$$\frac{64}{25 + h} + \frac{16}{h} = 1, \quad \text{or} \quad h^2 - 55h - 400 = 0.$$

Substituting in (2) gives the same equation for $k$. Hence, when we solve the quadratic, $h$ will be the positive root and $k$ the negative root. The student should do the work, and find

$$h = \frac{55 + 5\sqrt{185}}{2} \sim 61.5, \qquad k = \frac{55 - 5\sqrt{185}}{2} \sim -6.5.$$

The semiaxes of the ellipse are, approximately,

$$\sqrt{86.5} \sim 9.3 \quad \text{and} \quad \sqrt{61.5} \sim 7.8.$$

Since $\sqrt{18.5} \sim 4.3$ and $\sqrt{6.5} \sim 2.5$, the vertices of the hyperbola are approximately at $(\pm 4.3, 0)$, and the slopes of the asymptotes are approximately $\pm\frac{2.5}{4.3} = \pm 0.58$.

It was mentioned at the end of § 3-8 that when lines are drawn from the foci to a point $P$ on an ellipse, these lines make equal angles with the line tangent to the ellipse at $P$. Likewise, at the end of § 3-9 it was mentioned that when lines are drawn from the foci to a point $P$ on a hyperbola, the tangent at $P$ bisects the angle between these lines. From these two facts it follows that if an ellipse and a hyperbola have the same foci, then the two curves intersect at right angles. The student should satisfy himself on this matter by drawing a diagram. A different proof, by calculus and algebra, is indicated in Exercise 3(e).

**EXERCISES**

**1.** (a) Show that the ellipse and the hyperbola

$$\frac{x^2}{169} + \frac{y^2}{144} = 1, \qquad \frac{x^2}{9} - \frac{y^2}{16} = 1$$

are confocal.

(b) Find the values of $c^2$, $h$, $k$ to make these equations take the forms (1) and (2), respectively. Then find the first quadrant intersection of these curves. Check by drawing a figure.

(c) Find the slopes of these curves at their point of intersection, and verify that the curves cut at right angles.

**2.** Proceed as in Exercise 1 with the curves

$$\frac{x^2}{169} + \frac{y^2}{25} = 1, \qquad \frac{x^2}{108} - \frac{y^2}{36} = 1.$$

**3.** (a) Verify algebraically that, if $(x, y)$ is given subject to the single restriction that $|x| > c$ if $y = 0$, then there is a unique positive value of $h$ satisfying (1).

(b) Likewise verify that, if $(x, y)$ is given subject to the two restrictions that $x \neq 0$ and that $|x| < c$ if $y = 0$, then there is a unique negative value of $k$ satisfying (2).

(c) If $(x, y)$ is given with $x > 0$, $y > 0$, and if $h$ and $k$ are chosen so as to give, respectively, the ellipse and the hyperbola through $(x, y)$ [from (1) and (2)], show that $h + k = x^2 + y^2 - c^2$.

(d) In the situation of (c), express $x$ and $y$ in terms of $h$, $k$, and $c$.

(e) Calculate, in terms of $h$, $k$, and $c$, the slopes of the ellipse (1) and the hyperbola (2) at their point of intersection in the first quadrant. Note that these slopes are negative reciprocals of each other.

## 7-6 Translation and Rotation of Axes

For some purposes in analytic geometry it is useful to shift attention from one coordinate system to another. In this section we shall consider the change from one rectangular coordinate system to another. Such changes may be made in order to simplify the form of an equation which is being studied. Another reason for shifting attention from one coordinate system to another is evident in certain physical problems. If we are studying the motion of some solid object, we may wish to consider two coordinate systems: one system rigidly attached to the object, and one system which remains at rest. The system which is attached to the moving object is then being translated or rotated (or both) in relation to the system at rest.

*Translation of Axes*

Let us consider two rectangular coordinate systems, one with $x$ and $y$ axes, and one with $u$ and $v$ axes. Suppose the $uv$-system has the same orientation as the $xy$-system, with the $x$-axis parallel to the $u$-axis. Then we say that one system can be *translated* into the other (see Fig. 7-14). Let

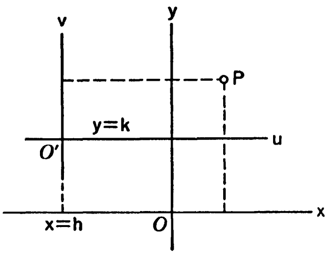



**Fig. 7-14**

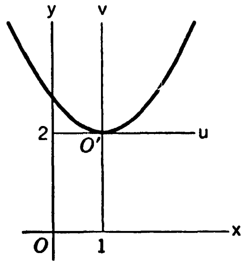



**Fig. 7-15**

the origin $O'$ of the $uv$-system be at $x = h$, $y = k$. Then, if a point $P$ has coordinates $(x, y)$ in one system and $(u, v)$ in the other, we see that

$$u = x - h, \qquad v = y - k, \tag{1}$$

**Example 1:** Find what the equation $3x^2 - 6x - 4y + 11 = 0$ becomes upon translation to new axes with origin at $x = 1$, $y = 2$.

Here we have $u = x - 1$, $v = y - 2$, or $x = u + 1$, $y = v + 2$. On substituting in the given equation, we have

$$3(u + 1)^2 - 6(u + 1) - 4(v + 2) + 11 = 0.$$

After simplification, this becomes

$$3u^2 - 4v = 0.$$

We recognize the equation as that of a parabola (see Fig. 7-15).

Sometimes it is left to us to discover a convenient choice for the location of the origin of the translated coordinate system.

**Example 2:** Make a translation in such a way as to get rid of the first-degree terms in the equation

$$xy - 2x + 3y = 10.$$

Here we use (1) with $h$ and $k$ left as literal constants at first. Our equation becomes

$$(u + h)(v + k) - 2(u + h) + 3(v + k) = 10.$$

Now multiply out and collect like terms. The result is

$$uv + (k - 2)u + (h + 3)v + hk - 2h + 3k = 10.$$

We wish to eliminate the first-degree terms in $u$ and $v$, so we now choose $k = 2$, $h = -3$. This gives

$$uv = 4,$$

which we recognize as the equation of a rectangular hyperbola with the $u$ and $v$ axes as asymptotes. See Fig. 7-16.

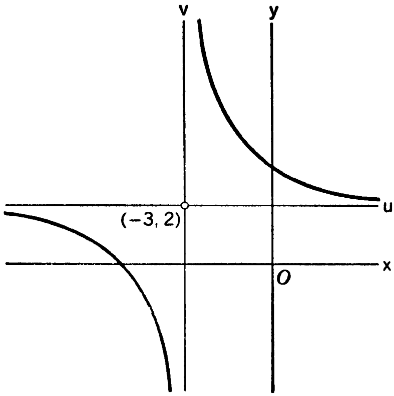



**Fig. 7-16**

*Rotation of Axes*

Suppose that the $xy$-system and the $uv$-system have the same origin and that the counterclockwise angle from the positive $x$-axis to the positive $u$-axis is $\phi$. Then the change from one system to the other is called a *rotation of axes*. In order to find the relations between the coordinates $(x, y)$ and $(u, v)$ for a given point $P$, we refer to Fig. 7-17. Let $\theta$ be the counter-clockwise angle from the positive $x$-axis to $OP$, and let $r$ denote the distance $OP$. Then

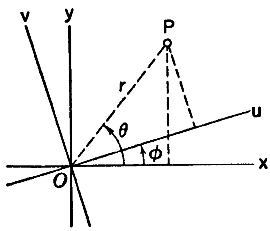



**Fig. 7-17**

$$\sin\theta = \frac{y}{r}, \qquad \cos\theta = \frac{x}{r},$$

$$\sin(\theta - \phi) = \frac{v}{r}, \qquad \cos(\theta - \phi) = \frac{u}{r}.$$

Hence $\quad u = r\cos(\theta - \phi) = r\cos\theta\cos\phi + r\sin\theta\sin\phi,$

or $u = x\cos\phi + y\sin\phi$. A formula for $v$ is found in the same way. So we have

$$\left.\begin{aligned} u &= x\cos\phi + y\sin\phi, \\ v &= -x\sin\phi + y\cos\phi. \end{aligned}\right\} \quad (2)$$

It is also desirable to express $x$ and $y$ in terms of $u$ and $v$. This can be done by solving (2) as simultaneous equations. But a more clever method is the following: Just as we pass from the $xy$-system to the $uv$-system by a counterclockwise rotation through an angle $\phi$, so we can pass from the $uv$-system to the $xy$-system by a further rotation through an angle $2\pi - \phi$ (look at Fig. 7-17). Hence we can exchange $u$ with $x$ and $v$ with $y$ in equations (2) if we also put $2\pi - \phi$ in place of $\phi$. Since $\sin(2\pi - \phi) = -\sin\phi$ and $\cos(2\pi - \phi) = \cos\phi$, we obtain

$$\left.\begin{aligned} x &= u\cos\phi - v\sin\phi, \\ y &= u\sin\phi + v\cos\phi. \end{aligned}\right\} \quad (3)$$

**Example 3:** Consider the equation

$$Ax^2 + 2Bxy + Ay^2 = 1, \quad (4)$$

where $A$ and $B$ are constants, not both zero (note that the coefficients of $x^2$ and $y^2$ are equal). We wish to know what can be said about the locus of this equation. Let us see what happens to the equation if we rotate the axes with $\phi = \pi/4$. For this case equations (3) become

$$x = \frac{u - v}{\sqrt{2}}, \qquad y = \frac{u + v}{\sqrt{2}}.$$

Then $x^2 + y^2 = u^2 + v^2$ and $xy = (u^2 - v^2)/2$, so that (4) becomes

$$(A + B)u^2 + (A - B)v^2 = 1. \quad (5)$$

From this form it is easy to decide the nature of the locus. It is an ellipse (or perhaps a circle) if $A + B$ and $A - B$ are both positive, and a hyperbola if $A + B$ and $A - B$ are of opposite signs. Further discussion of the different possibilities is considered in Exercise 8. As a particular case consider the equation

$$5x^2 - 6xy + 5y^2 = 32,$$

which can be written in the form (4) with $A = \frac{5}{32}$, $B = -\frac{3}{32}$. In the $uv$-system the equation becomes

$$\frac{u^2}{16} + \frac{v^2}{4} = 1.$$

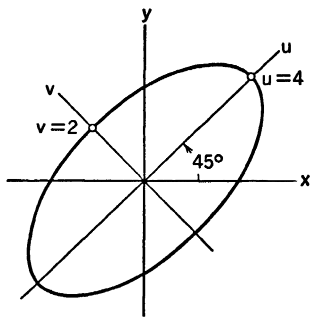



**Fig. 7-18**

Hence the locus is an ellipse with center at the origin, major axis 8, minor axis 4, and foci on the $u$-axis (see Fig. 7-18).

An alert student may wish to know how it could be predicted in advance that a rotation with $\phi = \pi/4$ would simplify equation (4) so that the locus could be identified. Would some other choice of $\phi$ do as well? We leave this question for speculation. It is related to a more general problem which we shall consider in § 7-7.

## EXERCISES

**1.** Make a translation of axes so as to get rid of the first-degree terms in the equation. Then, from the appearance of the equation in the *uv*-system, identify the locus and draw it. Show both sets of axes in your diagram.

(a) $9x^2 + 4y^2 + 18x - 16y = 11.$
(b) $4x^2 - 9y^2 - 16x + 18y = 29.$
(c) $xy + x + 6y = 0.$
(d) $x^2 - 4y^2 - 6x - 32y = 59.$
(e) $x^2 + y^2 + 6x - 14y = -22.$

**2.** Follow the instructions of Exercise 1.

(a) $9x^2 + 25y^2 + 18x - 50y = 191.$
(b) $xy + 2x - 5y = 18.$
(c) $9x^2 + 4y^2 - 15x - y = -2.$
(d) $4x^2 - y^2 - 16x - 6y = 0.$
(e) $xy + 2x - 3y = 8.$

**3.** Make a rotation with the indicated angle and find the new form of the equation. Identify the locus and draw it, showing both sets of axes.

(a) $x^2 - \sqrt{3}xy + 2y^2 = 10, \quad \phi = \pi/6.$
(b) $4x^2 + 3\sqrt{3}xy + y^2 = 22, \quad \phi = 120°.$
(c) $x^2 - \sqrt{3}xy + 12 = 0, \quad \phi = \pi/3.$

**4.** Identify the locus of:

(a) $17x^2 - 16xy + 17y^2 = 225.$
(b) $3x^2 - 10xy + 3y^2 = 32.$
(c) $x^2 + xy + y^2 + 1 = 0.$
(d) $3x^2 - 6xy + 3y^2 + 8 = 0.$
(e) $5x^2 + 10xy + 5y^2 = 16.$

**5.** Make the rotation with $\sin\phi = \frac{3}{5}$, $\cos\phi = \frac{4}{5}$ and find the new form of the equation $52x^2 - 72xy + 73y^2 = 100$. Hence identify the locus.

**6.** Make the rotation with $\phi = \tan^{-1}\frac{2}{3}$ and find the new form of the equation $5x^2 + 24xy - 5y^2 = -325$. Hence identify the locus.

**7.** What does the equation $xy = -16$ become after a rotation with $\phi = 135°$?

**8.** Use equation (5) to complete the following classification of possible loci represented by equation (4):

| $A + B$ | $A - B$ | Nature of Locus |
|---|---|---|
| + | + | $B \neq 0$, ellipse; $B = 0$, circle |
| + | − | |
| − | + | |
| − | − | |
| 0 | + | |
| + | 0 | |
| 0 | − | |
| − | 0 | |

**9.** Under what circumstances will (4) represent (a) a circle? (b) a rectangular hyperbola? (c) two parallel lines?

**10.** (a) If we make the translation of axes (1), show that the equation $y = \sin x$ takes the form $v = A \sin u + B \cos u + C$, where $A$, $B$, $C$ are certain constants. State exactly how $A$, $B$, $C$ are expressed in terms of $h$ and $k$, and observe that $A^2 + B^2 = 1$.

(b) Show that the translation $x + (\pi/4) = u$, $y + 1 = v$ makes the equation $y = (1/\sqrt{2})(\sin x + \cos x) - 1$ take the form $v = \sin u$. Sketch the graph, showing both old and new axes.

(c) What is the appropriate translation to reduce $y = (1/\sqrt{2})(\sin x - \cos x) + 2$ to the form $v = \sin u$?

(d) The same question as in (c) for $y = \frac{1}{2} \sin x + (\sqrt{3}/2) \cos x + 1$; for $y = -(\sqrt{3}/2) \sin x - \frac{1}{2} \cos x + 4$.

(e) If $a^2 + b^2 = 1$, explain how to make the equation $y = a \sin x + b \cos x + c$ take the form $v = \sin u$ by a translation of axes. What form can be achieved in case $0 < a^2 + b^2 \neq 1$?

## 7-7 Homogeneous Quadratic Forms

An expression of the type

$$Ax^2 + 2Bxy + Cy^2 \tag{1}$$

is called a *homogeneous quadratic form* in $x$ and $y$. Such forms occur in many situations and in various contexts in mathematics. We have not the space in this book to deal with the ways in which quadratic forms are important in connection with the kinetic energy of mechanical systems, or in connection with certain ideas in the theory of probability. We shall study the form (1) in relation to the problem of identifying the locus of an equation of the form

$$Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F = 0. \tag{2}$$

The systematic classification of results relating to (2) will be given in § 7-8. For the present we concentrate attention on the following problem: What important facts can be observed in connection with the changes made in the quadratic form (1) by rotation of axes?

The first thing we observe is this: If we make *any* rotation of the axes, $x$ and $y$ are expressions of the first degree in $u$ and $v$, and hence (1) is changed into a homogeneous quadratic form in $u$ and $v$. The coefficients in the new form can be computed in terms of $A$, $B$, $C$ and $\sin \phi$, $\cos \phi$. The procedure is to compute $x^2$, $y^2$ and $xy$ from (3) in § 7-6, and substitute in (1). The result is

$$au^2 + 2buv + cv^2, \tag{3}$$

where

$$a = A \cos^2 \phi + 2B \sin \phi \cos \phi + C \sin^2 \phi \tag{4}$$

$$2b = 2B \cos 2\phi - (A - C) \sin 2\phi \tag{5}$$

$$c = A \sin^2 \phi - 2B \sin \phi \cos \phi + C \cos^2 \phi. \tag{6}$$

In obtaining equation (5) we use the formulas $\cos 2\phi = \cos^2 \phi - \sin^2 \phi$, $\sin 2\phi = 2 \sin \phi \cos \phi$.

Now suppose that $B \neq 0$. From (5) we see that $b = 0$ if we choose $\phi$ so that

$$\operatorname{ctn} 2\phi = \frac{A - C}{2B}. \tag{7}$$

Since this choice can always be made, we have proved the following assertion:

*Given the quadratic form* (1) *in which* $B \neq 0$, *it is possible to make a rotation of the axes in such a way that in the new coordinates the uv-term is eliminated and the quadratic form has the appearance*

$$au^2 + cv^2. \tag{8}$$

For the actual work of carrying out the rotation of the axes with an angle $\phi$ determined by (7), the following procedure is convenient: Choose $2\phi$ as an angle between 0 and $\pi$ with cotangent given by (7); then $0 < \phi < \pi/2$, so that $\sin \phi$ and $\cos \phi$ are positive. Then, from known trigonometric identities,

$$\sin \phi = \left(\frac{1 - \cos 2\phi}{2}\right)^{1/2}, \qquad \cos \phi = \left(\frac{1 + \cos 2\phi}{2}\right)^{1/2}. \tag{9}$$

Knowing $\operatorname{ctn} 2\phi$, we compute $\cos 2\phi$, and then we compute $\sin \phi$ and $\cos \phi$. After this we can write the equations for making the rotation of axes. Except in rather special cases the numerical work in all this is rather awkward.

**Example 1:** Simplify the equation $8x^2 + 4xy + 5y^2 = 36$ by rotation of axes.

Here we have

$$\operatorname{ctn} 2\phi = \frac{8 - 5}{4} = \frac{3}{4}, \qquad \cos 2\phi = \frac{3}{5},$$

$$\sin \phi = \frac{1}{\sqrt{5}}, \qquad \cos \phi = \frac{2}{\sqrt{5}}.$$

When we make the rotation of axes by equations (3) from § 7-6, we find that our original equation becomes

$$9u^2 + 4v^2 = 36.$$

It represents an ellipse with foci on the $v$-axis.

When $b = 0$ in (3), there is a procedure for finding the values of $a$ and $c$ without having to compute $\sin \phi$ and $\cos \phi$. This procedure furnishes us with two numbers, one of which is $a$ and one of which is $c$, but it does not tell us which is which. For example, it might tell us that $a$ and $c$ are either 9 and 4 or 4 and 9. This lack of certainty is inevitable when we do not know $\phi$; for, if a certain choice of $\phi$ makes $b = 0$, an increase of $\phi$ by $\pi/2$ would also make $b = 0$, and this would have the effect of exchanging the values of $a$ and $c$. This can be seen by putting $\phi + (\pi/2)$ in place of $\phi$ in equations (4)–(6). This method for finding $a$ and $c$ is as follows:

*When the quadratic form* (1) *is changed to the form* (8) *by a rotation of axes, the coefficients* $a$, $c$ *are the roots of the quadratic equation*

$$t^2 - (A + C)t + (AC - B^2) = 0. \tag{10}$$

We remark that this equation can be written in the following way, using a determinant:

$$\begin{vmatrix} A - t & B \\ B & C - t \end{vmatrix} = 0. \tag{11}$$

The proof of the assertion just made depends upon two important facts about what happens to the quadratic form (1) when we make an *arbitrary* rotation of the axes. These facts are that

$$a + c = A + C, \tag{12}$$

and

$$b^2 - ac = B^2 - AC. \tag{13}$$

The truth of (12) can be seen at once by adding (4) and (6). It is a more tedious job to prove (13) by using (4), (5), and (6), but all that is involved is simple calculation and the use of trigonometric identities. We leave the details as an exercise for the student.

Once (12) and (13) are known to be true, let us proceed to the assertion made about the roots of (10). We see that

$$(t - a)(t - c) = t^2 - (a + c)t + ac.$$

Now, if $b = 0$, we see from (12) and (13) that

$$(t - a)(t - c) = t^2 - (A + C)t + (AC - B^2).$$

Hence the roots of (10) are the same as the roots of $(t - a)(t - c) = 0$, namely $a$ and $c$.

**Example 2:** Identify the curve

$$7x^2 - 8xy + y^2 = 9$$

without explicitly making a rotation of axes.

In this case $A = 7$, $B = -4$, $C = 1$, so that equation (10) becomes

$$t^2 - 8t - 9 = 0,$$

with roots $t = -1, 9$. Hence we know that a certain rotation of axes will bring our equation to the form

$$9u^2 - v^2 = 9.$$

The curve is therefore a hyperbola.

## EXERCISES

1. Without explicitly making a rotation, identify each locus. If it is an ellipse, give the length of its axes. If it is a hyperbola, give the distance

between its vertices. If it is a pair of parallel lines, give the distance between them.

(a) $2x^2 - 4xy + 5y^2 = 36$. (d) $25x^2 - 120xy + 144y^2 = -1$.
(b) $9x^2 + 24xy + 2y^2 = 126$. (e) $9x^2 - 42xy + 49y^2 = 0$.
(c) $9x^2 + 24xy + 16y^2 = 100$. (f) $2x^2 + 5xy + 2y^2 = 8$.

**2.** Follow the instructions of Exercise 1.

(a) $x^2 - 4xy - 2y^2 = 24$. (d) $3x^2 - 2xy - 3y^2 + \sqrt{10} = 0$.
(b) $5x^2 + 4xy = 16$. (e) $73x^2 + 72xy + 52y^2 = 100$.
(c) $4x^2 - 4xy + y^2 = 5$. (f) $x^2 - 3xy + y^2 + 8 = 0$.

**3.** Carry out a rotation of axes which will change each quadratic form into one of the form $au^2 + cv^2$. Then identify the locus. If it is an ellipse or a hyperbola, give the slope of the line through its foci. If the locus is a pair of parallel lines, give their slope and the distance between them.

(a) $9x^2 - 6xy + y^2 = 10$. (d) $\sqrt{3}x^2 - 3xy = 6$.
(b) $4x^2 + 24xy + 11y^2 = -80$. (e) $20x^2 - 12xy + 25y^2 = 16$.
(c) $25x^2 - 24xy + 32y^2 = 64$. (f) $4x^2 + 5xy - 8y^2 = -8$.

## 7-8 Equations of the Second Degree

Suppose we are confronted by an equation of the form

$$Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F = 0. \tag{1}$$

It is our purpose in this section to show that we can always follow procedures which will enable us to identify the locus of this equation and to tell the position of the locus in relation to the coordinate axes. These procedures will in general involve both translations and rotations of axes, though in particular cases both things may not be necessary. Moreover, it is possible to identify the type of the locus, though not its location relative to the axes, without making any rotations or translations. There are three basic types: elliptic, hyperbolic, and parabolic.

It is the $xy$-term in (1) which causes the real difficulty in identifying the locus. Let us first consider what can be said about the locus when $B = 0$. We put aside the case in which $A$, $B$, and $C$ are *all* zero, for then the equation is linear and the locus is a straight line. There are then, three essentially different cases to consider when $B = 0$:

I. If $A \neq 0$, $C \neq 0$, and $A$ and $C$ are of the same sign, the situation is like that considered in equation (11) at the end of § 3-8. The locus is, generally speaking, an ellipse, but it may in particular be a circle or a point, or there may be no locus at all. The center of the ellipse may be located by completion of squares, and the equation may then be simplified by a translation of axes.

II. If $A \neq 0$, $C \neq 0$ and $A$ and $C$ are of opposite signs, the locus is, generally speaking, a hyperbola, but it may in particular cases be two intersecting lines. The center of the hyperbola may be located by completion of squares. This sort of thing was considered in § 3-9.

III. If either $A = 0$ or $C = 0$ (but not both), the locus is, generally speaking, a parabola, though it may in particular cases be one line, or two parallel lines, or there may be no locus. We illustrate in the following example.

**Example 1:** Consider the equation

$$3x^2 - 6x + 8py + q = 0,$$

where $p$ and $q$ are unspecified constants. Completing the square in $x$, we have

$$x^2 - 2x + 1 = -\frac{8}{3}py + 1 - \frac{q}{3}.$$

If $p = 0$, this becomes

$$(x - 1)^2 = \frac{3 - q}{3},$$

the locus is two parallel lines if $q < 3$, one line if $q = 3$, and no locus if $q > 3$. If $p \neq 0$, the equation becomes

$$(x - 1)^2 = -\frac{8}{3}p\left(y - \frac{3 - q}{8p}\right).$$

The locus is a parabola with vertex at $\left(1, \frac{3 - q}{8p}\right)$. We may, if we wish, translate the axes to make this point the new origin.

Let us now assume that $B \neq 0$ in (1). In this case we can make a rotation of axes, as described in § 7-7, so as to change the first three terms of (1) into an expression $au^2 + cv^2$. The last three terms of (1) will be changed into an expression $du + ev + f$, so that equation (1) will become

$$au^2 + cv^2 + du + ev + f = 0. \tag{2}$$

For this equation we have the three cases previously described. I: $a$ and $c$ of the same sign, i.e., $ac > 0$; II: $a$ and $c$ of opposite signs, i.e., $ac < 0$; III: $a = 0$ or $c = 0$, i.e., $ac = 0$. But, by equation (13) in § 7-7, $B^2 - AC = b^2 - ac$, so in the present case $B^2 - AC = -ac$. Hence we can make the following assertions about the locus of (1) *without* actually making any rotations or translations:

I. If $B^2 - AC < 0$, the locus is an ellipse, a circle, or a point, or there is no locus.

II. If $B^2 - AC > 0$, the locus is a hyperbola or two intersecting lines.

III. If $B^2 - AC = 0$, the locus is a parabola, two parallel lines, one line, or there is no locus.

If one actually wishes to carry out the work of making the rotation of axes, it is best to make a translation first when $B^2 - AC \neq 0$, so as to get the center of the ellipse or hyperbola at the new origin. This is done by choosing the translation so as to eliminate the linear terms from the equation. We illustrate with an example.

**Example 2:** Consider the equation

$$7x^2 - 8xy + y^2 - 50x + 26y + 79 = 0. \tag{3}$$

Here $B^2 - AC = 16 - 7 > 0$, so we have the hyperbolic case. We make the translation $x = u + h$, $y = v + k$, with $h$ and $k$ to be determined. When we substitute and simplify, the coefficients of $u$ and $v$ are found to be

$$14h - 8k - 50 \quad \text{and} \quad -8h + 2k + 26,$$

respectively. We set these expressions equal to zero and solve for $h$ and $k$:

$$\begin{aligned} 14h - 8k &= 50, \\ -8h + 2k &= -26, \qquad h = 3, \quad k = -1. \end{aligned}$$

Details of the algebra are left to the student. When we use these values of $h$ and $k$, the new form of (3) with the linear terms in $u$ and $v$ missing, is

$$7u^2 - 8uv + v^2 = 9.$$

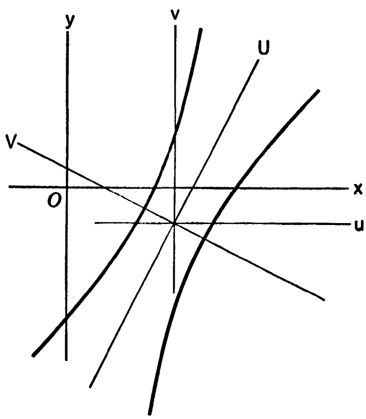



**Fig. 7-19**

Except for the change in letters, this is the same as the equation of Example 2, § 7-7. If we make the rotation of axes with

$$\operatorname{ctn} 2\phi = \frac{7 - 1}{-8} = -\frac{3}{4}, \qquad 0 < \phi < \frac{\pi}{2},$$

we find that the equation takes the form

$$-U^2 + 9V^2 = 9,$$

where $U$ and $V$ are the final coordinates. So the locus is a hyperbola with center at $x = 3$, $y = -1$ and foci on the line $U = 0$. The situation is shown in Fig. 7-19.

## EXERCISES

**1.** Identify the type of each of the following equations. Then simplify the equation by rotations or translations, or both. Draw a figure showing the locus and all the coordinate axes which are used.

(a) $3x^2 - 4xy + 8x - 1 = 0$.
(b) $16x^2 - 24xy + 9y^2 - 60x - 80y + 400 = 0$.
(c) $5x^2 + 6xy + 5y^2 + 22x - 6y + 21 = 0$.
(d) $9x^2 - 6xy + y^2 + 6\sqrt{10}\,(3x - y) + 50 = 0$.
(e) $17x^2 - 12xy + 8y^2 - 68x + 24y - 12 = 0$.
(f) $5x^2 + 4xy - y^2 + 24x - 6y - 5 = 0$.

**2.** Proceed as directed in Exercise 1.

(a) $144x^2 - 120xy + 25y^2 - 29x - 2y - 1 = 0$.
(b) $24xy - 7y^2 - 120y - 144 = 0$.
(c) $25x^2 + 36xy + 40y^2 - 308x - 384y - 108 = 0$.
(d) $x^2 + 2xy + 3y^2 - 4x - 8y + 6 = 0$.
(e) $18x^2 + 24xy + 8y^2 - 21x - 14y + 3 = 0$.
(f) $36x^2 - 96xy + 64y^2 - 360x + 480y + 675 = 0$.

**3.** Show that there is just one parabola tangent to the $x$-axis at $(4, 0)$ and tangent to the line $y = x$ at $(3, 3)$. Find the equation of it in the form (1). Find the slope of the axis of the parabola. If you seem to be getting *two* parabolas, examine carefully the locus of the second equation and explain its geometric relation to the given points.

**4.** A line of slope $m$ is drawn through $(2, 0)$ intersecting the line $2y = x$ at $A$ and the line $y = 2x$ at $B$. If $P$ is the mid-point of $AB$, express the coordinates $(x, y)$ of $P$ in terms of $m$. Then, treating $m$ as a parameter, consider the locus of $P$. By expressing $m$ in terms of $x$ and $y$ and then eliminating $m$, show that $x$ and $y$ satisfy a certain equation of the second degree. Identify the locus and draw it.

*CHAPTER* VIII

# LOGARITHMIC AND EXPONENTIAL FUNCTIONS

## 8-1 Exponents and Logarithms

Students in high school become familiar with the use of exponents, and they learn about logarithms as defined in terms of exponents. Let us briefly review the facts and definitions as they appear in this customary approach.

For our purposes here we shall consider expressions of the type $a^u$, where $a > 0$ and $u$ is any real number. There is no difficulty in explaining exactly what $a^u$ means if $u$ is an integer, and we assume the student knows these explanations. We have presented some discussion of fractional exponents in § 3-6; we recapitulate in brief here. If $p$ and $q$ are integers, with $q > 0$, the definition of $p/q$ as an exponent is as follows:

$$a^{p/q} = (\sqrt[q]{a})^p,$$

where $\sqrt[q]{a}$ is the unique *positive* number whose $q$th power is $a$. In order to be assured that there is in fact exactly one positive number whose $q$th power is the given positive number $a$, it suffices to know that $x^q$ is a continuous function of $x$ which increases as $x$ increases, and that $x^q \to +\infty$ when $x \to +\infty$. Then $x^q$ must pass through the value $a$, by Theorem 6-A.

If $u$ is an irrational number, i.e., one which cannot be represented as the quotient of two integers, the definition of $a^u$ is a somewhat complicated matter for the students at a very elementary level. One natural way to make the definition depends upon the fact that an irrational number can be approximated as closely as one wishes by rational numbers. For instance, one may think of the irrational number as a nonterminating decimal; it can then be approximated by terminating decimals with more and more decimal places. These latter decimals will be rational numbers. Suppose, then, that $u_1, u_2, \cdots$ are rational numbers such that $u_n \to u$ as $n \to \infty$. The meaning of $a^{u_n}$ is known; and it turns out that $a^{u_n}$ approaches a limiting value as $n$ increases. This limiting value may be defined to be the value of $a^u$. This method of defining $a^u$ when $u$ is irrational is logically satisfactory, but a considerable amount of time and care must be spent in detailed verification that everything actually works out the way one hopes and expects. We pass over these details and merely state that, as a final result, the laws of exponents hold in the following form for $a > 0$, $b > 0$ and all real values of $u$ and $v$, both rational and irrational:

$$a^u a^v = a^{u+v}, \tag{1}$$

$$(a^u)^v = a^{uv}, \tag{2}$$

$$(ab)^u = a^u b^u. \tag{3}$$

Next, it is natural to consider $a^x$ as a function of $x$ and investigate its properties. Here again we state the essential facts without going into the logical details of how the facts are established. We assume $a > 0$ and $a \neq 1$. The case $a = 1$ is dismissed, since $1^x = 1$ for all $x$. For definiteness let us assume $a > 1$. Then $a^x$ is a continuous function of $x$ which increases as $x$ increases; the values of $a^x$ are all positive, and

$$\lim_{x \to -\infty} a^x = 0, \qquad \lim_{x \to +\infty} a^x = +\infty. \tag{4}$$

To construct the graph, plot the points corresponding to several integral values of $x$, both positive and negative. The curve can then be filled in smoothly. See Fig. 8-1. If $0 < a < 1$, $a^x$ *decreases* as $x$ increases, and the graph has the appearance shown in Fig. 8-2.

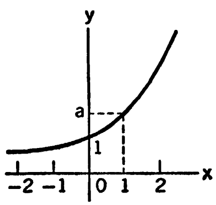



**Fig. 8-1**

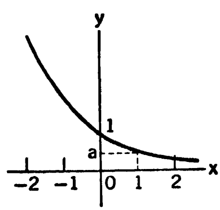



**Fig. 8-2**

Once this much about exponents is known or taken for granted, it is rather easy to define logarithms and develop some of their properties. The properties of $a^x$ show clearly that if $a > 0$ and $a \neq 1$, to each positive $y$ corresponds a unique $x$ such that $a^x = y$. This $x$ is called the logarithm to the base $a$ of $y$, and we write $x = \log_a y$. The laws of exponents become properties of logarithms. If $a^u = A$ and $a^v = B$, then $AB = a^{u+v}$. Hence, since $u = \log_a A$, and so on, we see that

$$\log_a (AB) = \log_a A + \log_a B. \tag{5}$$

The law of exponents in (2) leads to the following law of logarithms:

$$\log_a A^v = v \log_a A. \tag{6}$$

Especially to be noted are the particular facts,

$$\log_a a = 1 \quad \text{and} \quad \log_a 1 = 0. \tag{7}$$

If we wish to study the function $f(x) = \log_a x$, we note that $y = \log_a x$ is equivalent to $a^y = x$. In particular, $x$ must be positive in order that $\log_a x$ may be defined. The appearance of the graph of $y = \log_a x$ can be deduced from the appearance of the graph of $y = a^x$; we must exchange the roles played by $x$ and $y$. Figure 8-3 shows the graph of $y = \log_a x$ when $a > 1$. In this case $\log_a x$ increases continuously as $x$ increases. The facts corresponding to those in (4) are

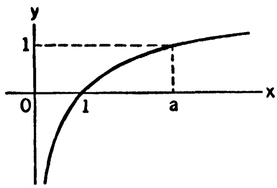


**Fig. 8-3**

$$\lim_{x \to 0^+} \log_a x = -\infty, \qquad \lim_{x \to +\infty} \log_a x = +\infty. \tag{8}$$

The conception of logarithms was developed near the end of the 16th century. Their earliest use seems to have been mainly for simplifying computations in astronomy; a number of tables were constructed early in the 17th century. The computational usefulness of logarithms (as in high school trigonometry, for instance) is only one aspect of their importance for mathematics. Actually, the logarithm as a function is of very great importance in theoretical work. It is our immediate objective in the next few sections to learn about logarithmic and exponential functions in connection with differentiation and integration.

## EXERCISES

**1.** Find the value of each logarithm.

(a) $\log_2 32$.
(b) $\log_{1/2} 64$.
(c) $\log_3 \dfrac{1}{81}$.
(d) $\log_{9/4} \frac{2}{3}$.
(e) $\log_{0.1} 10{,}000$.
(f) $\log_9 27$.

**2.** Deduce from (5) that $\log_a \dfrac{A}{B} = \log_a A - \log_a B$; then from this show that $\log_a B^{-1} = -\log_a B$.

**3.** If $0 < a < 1$ and $b = a^{-1}$, show that $\log_a x = -\log_b x$.

**4.** If $\log_a x = \log_a y$, it follows that $x = y$. Why? Use this fact to show that $b^y = a^{y \log_a b}$. If this method is used to express $(ab)^u$ as a power of $a$, one may then use (5), (7), and (1) to deduce (3). Do this.

**5.** Show that $\log_a x = (\log_a b)(\log_b x)$. *Suggestion*: Set $x = b^y$ and use the first part of Exercise 4. Show also that $\log_a b = (\log_b a)^{-1}$.

**6.** Explain why $\log_a (a^x) = x$ and why $b^{\log_b y} = y$.

**7.** Show that, to any base $a$,

(a) $2 \log_a \sin\theta = \log_a (1 - \cos\theta) + \log_a (1 + \cos\theta)$ if $0 < \theta < \pi$;

(b) $2 \log_a \cos\frac{\theta}{2} = \log_a \frac{1 + \cos\theta}{2}$ if $-\pi < \theta < \pi$;

(c) $\log_a \tan\frac{\theta}{2} = \log_a \sin\theta - \log_a (1 + \cos\theta)$ if $0 < \theta < \pi$.

**8.** If $f(x) = \log_a x$, show that

$$\frac{f(x+h) - f(x)}{h} = f\left[\left(1 + \frac{h}{x}\right)^{1/h}\right].$$

**9.** Show that

$$\log_a (x + \sqrt{x^2 - 1}) = -\log_a (x - \sqrt{x^2 - 1}),$$

and that

$$\log_a (\csc x - \operatorname{ctn} x) = -\log_a (\csc x + \operatorname{ctn} x).$$

**10.** If $y = \log_a (x + \sqrt{x^2 + 1})$, show that $x = \frac{1}{2}(a^y - a^{-y})$.

## 8-2 A New Approach

The basic property of logarithms expressed in (5) of § 8-1 can be written in the form

$$f(AB) = f(A) + f(B), \tag{1}$$

where $f(x) = \log_a x$. Let us consider for a moment *any* function $f$ that obeys the equation (1) and is such that $f$ is defined when $x > 0$. In the immediately following reasoning we dismiss logarithms from our minds and focus all our attention on the functional notation. First of all, putting $A = B = 1$ in (1), we see that $f(1) = f(1) + f(1)$, and hence

$$f(1) = 0. \tag{2}$$

Next, putting $B = A^{-1}$ in (1), we see that $f(A) + f(A^{-1}) = f(AA^{-1}) = f(1) = 0$, or

$$f(A^{-1}) = -f(A). \tag{3}$$

Hence $f(A/B) = f(AB^{-1}) = f(A) + f(B^{-1})$, or

$$f\left(\frac{A}{B}\right) = f(A) - f(B). \tag{4}$$

Now let us attempt to use (1) to find the derivative $f'(x)$. By definition,

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

if the limit exists. If $x$ and $x + h$ are positive, let $t = h/x$. Then, by (4),

$$f(x + h) - f(x) = f\left(\frac{x + h}{x}\right) = f(1 + t),$$

and so, in view of (2),

$$\frac{f(x + h) - f(x)}{h} = \frac{f(1 + t)}{tx} = \frac{1}{x}\frac{f(1 + t) - f(1)}{t}. \tag{5}$$

Since $h \to 0$ is equivalent to $t \to 0$, we see from (5) that

$$f'(x) = \frac{f'(1)}{x}, \tag{6}$$

provided that the limit defining $f'(1)$ exists.

Of course, we do not yet know anything about $f'(1)$. However, let us now go back to the fact that $f(x) = \log_a x$ and that (since $f(1) = 0$)

$$f'(1) = \lim_{t \to 0} \frac{f(1 + t)}{t} = \lim_{t \to 0} \frac{1}{t} \log_a (1 + t).$$

We see that

$$f'(1) = \lim_{t \to 0} \log_a (1 + t)^{1/t}.$$

This brings us up against the problem of finding the value of the limit

$$\lim_{t \to 0} (1 + t)^{1/t}. \tag{7}$$

This is not an easy problem, when approached directly. We shall approach the whole matter in a different way. Ultimately we shall be able to show that the limit in (7) does exist; the limit is a certain irrational number whose decimal form to three places is 2.718.

Our new approach has several advantages. Not only do we avoid the difficulties of dealing directly with (7), but we also eliminate the need for dealing with the logical details of the definition of $a^u$ for irrational values of $u$ by the pattern referred to in § 8-1.

The motivation for the new approach lies in (6). From $f'(t) = f'(1)/t$, $f(1) = 0$, and Theorem 6-D we infer that

$$\int_1^x \frac{f'(1)}{t}\, dt = f(x) - f(1) = f(x).$$

This gives a wholly new way of studying logarithms by using integrals. The value of $f'(1)$ turns out to be related to the base $a$. The simplest choice of base is that which makes $f'(1) = 1$.

We shall now start directly with the definition of a function $L$ by the formula

$$L(x) = \int_1^x \frac{dt}{t}, \qquad x > 0. \tag{8}$$

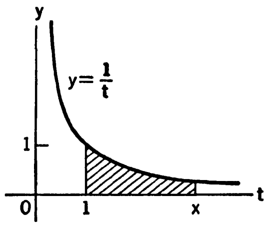



Fig. 8-4

Our deductions, starting with this definition, are logically independent of what has already been said about logarithms. The letter $L$ is used in (8) because $L(x)$ will turn out to be a logarithm.

We can think of $L(x)$ in terms of area under the curve $y = 1/t$ (see Fig. 8-4). Clearly

$$L(1) = 0. \tag{9}$$

We know by Theorem 6-C that

$$L'(x) = \frac{1}{x}. \tag{10}$$

The fact that $L'(x) > 0$ shows that $L(x)$ increases as $x$ increases. Also $L(x) > 0$ if $x > 1$ and $L(x) < 0$ if $0 < x < 1$.

Now consider $L(ax)$, where $a$ is a positive constant. Letting $u = ax$, we see by the chain rule (Theorem 3-E) that

$$\frac{d}{dx} L(u) = \frac{d}{du} L(u) \frac{du}{dx} = L'(u) \cdot a.$$

In view of (10), this becomes

$$\frac{d}{dx} L(ax) = \frac{1}{ax} \cdot a = \frac{1}{x}.$$

Since $L(x)$ and $L(ax)$ have the same derivative, they differ by a constant:

$$L(ax) - L(x) = C.$$

To find the value of $C$, put $x = 1$. Then, since $L(1) = 0$, we have $L(a) = C$. Therefore $L(ax) = L(x) + L(a)$. Changing the notation a bit, we write this in the form

$$L(AB) = L(A) + L(B). \tag{11}$$

We can use (11) to help us find out how $L(x)$ behaves as $x \to +\infty$ or as $x \to 0^+$. Putting $A = B$ in (11) we see that $L(A^2) = 2L(A)$. Then, putting $B = A^2$, we see that $L(A^3) = L(A) + L(A^2) = 3L(A)$. Proceeding by induction, we see that

$$L(A^n) = nL(A) \tag{12}$$

for each positive integer $n$. It is also easy to show that (12) holds when $n$ is a negative integer (see Exercise 1). Now, if we take $A = 2$ and note that $L(2) > 0$, we see from (12) that $L(2^n) = nL(2) \to +\infty$ as the positive integer $n$ increases. Likewise $L(2^{-n}) = -nL(2) \to -\infty$ as $n$ increases.

Since $2^n$ becomes large and $2^{-n}$ becomes small as $n$ increases, we see (from the fact that $L(x)$ increases as $x$ increases) that

$$\lim_{x \to +\infty} L(x) = +\infty \quad \text{and} \quad \lim_{x \to 0^+} L(x) = -\infty. \tag{13}$$

We now have enough information to form a pretty good notion of the appearance of the graph of $L(x)$. See Fig. 8-5. Since $L(x)$ is continuous, and increases when $x$ increases, there is a unique value of $x$ for which $L(x) = 1$.

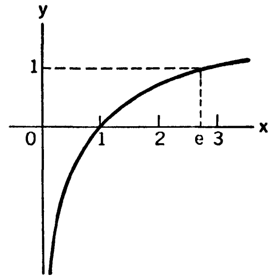



**Fig. 8-5**

DEFINITION. The letter $e$ is used to denote the unique positive number such that $L(e) = 1$.

It is not difficult to show that $L(\frac{5}{2}) < 1$ and $L(3) > 1$ (see Exercises 2, 3). Hence $2.5 < e < 3.0$. Later on in this book more exact estimates of $e$ can be made with ease. To six significant figures $e = 2.71828$. As with $\pi$, the decimal for $e$ is nonterminating and nonrepeating.

It is easy to show that (12) remains valid if $n$ is replaced by a fraction. Suppose $p$ and $q$ are integers, with $q > 0$. If $A > 0$, let $x = A^{1/q}$, so that $x^q = A$ and $x^p = A^{p/q}$. Then, using (12) for the case when $n$ is an integer, we have $L(A) = L(x^q) = qL(x)$, or $L(x) = (1/q)L(A)$. Also, $L(A^{p/q}) = L(x^p) = pL(x)$, whence

$$L(A^{p/q}) = \frac{p}{q} L(A). \tag{14}$$

We do not at this point go further and replace $p/q$ by an irrational exponent, for we are taking the point of view that an irrational power of $A$ is not yet satisfactorily defined. Such powers of $A$ will be considered in the next section.

**EXERCISES**

1. Prove that $L(A^{-n}) = -nL(A)$ if $n$ is a positive integer. Begin by considering the case $n = 1$.

2. By considering the areas of certain trapezoids, show that $L(\frac{5}{2}) < \frac{39}{40}$. See Fig. 8-4.

3. By considering the areas of certain trapezoids, show that $L(3) > \frac{16}{15}$. Use Fig. 8-4 and consider the tangents to the curve at $t = \frac{3}{2}, \frac{5}{2}$.

## 8-3 The New Method of Defining Powers

In this section we continue the logical development based on the definition of $L(x)$ by (8) in § 8-2. An inspection of the graph of $L(x)$ shows that each real number, no matter whether negative, zero, or positive, occurs as a value of $L(x)$ for a unique positive value of $x$. We now use this fact to define a new function, which we denote by $E$. By definition $E(x) = y$ if $y$ is the unique positive number for which $L(y) = x$. Thus $E(x)$ is defined for all values of $x$, and $E(x) > 0$. For convenience of reference we list the definition again:

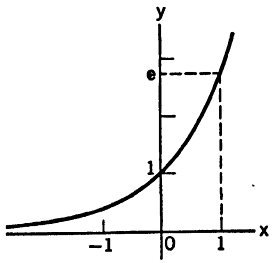


**Fig. 8-6**

$$E(x) = y \quad \text{means} \quad L(y) = x. \tag{1}$$

The function $E$ is the inverse of $L$ in the same sense that the inverse-sine function is the inverse of the sine function (see § 4-4).

The graph of $y = E(x)$ can be obtained from the graph of the function $L$. See Fig. 8-6 and refer to Fig. 8-5. Note that $E(1) = e$.

From the property of $L$ expressed in (11) of § 8-2 we can at once infer that

$$E(u + v) = E(u)E(v); \tag{2}$$

see Exercise 4.

The function $E$ is differentiable, since $L$ is, and since $L'(y) = 1/y \neq 0$. To calculate $E'(x)$, we start with $L(y) = x$ and differentiate with respect to $x$. By the chain rule and (10) in § 8-2, we have

$$1 = \frac{d}{dx} L(y) = \frac{d}{dy} L(y) \frac{dy}{dx} = \frac{1}{y} \frac{dy}{dx},$$

or

$$\frac{dy}{dx} = y = E(x).$$

Hence

$$E'(x) = E(x). \tag{3}$$

This is a very important formula.

From (1) we see that

$$E[L(y)] = y \quad \text{if} \quad y > 0 \tag{4}$$

and that

$$L[E(x)] = x \quad \text{for each } x. \tag{5}$$

From (4) we get a new method of expressing powers of a number. Suppose $a > 0$, and let $n$ be an integer. Then $a^n = E[L(a^n)]$, by (4). By (12) in § 8-2 we can then write

$$a^n = E[nL(a)].$$

This same formula, with $n$ replaced by a fraction, can be obtained by using § 8-2, (13). This points the way to an appropriate *definition* of $a^u$ when $u$ is irrational. The definition is given in the formula

$$a^u = E[uL(a)]. \tag{6}$$

Thus $a^u$ can *in all cases* be expressed by (6), in terms of the functions $E$ and $L$, whose properties we have been discussing.

From the known properties of $E$ and $L$ the following facts are clear: $a^u > 0$, and if $a > 1$, then $a^u$ increases as $u$ increases. (Why is $a > 1$ necessary for this?) The exponent law $a^{u+v} = a^u a^v$ is a consequence of (6) and (2); this should be verified by the student.

We can now make contact between the function $L$ and the usual definition of logarithms. Suppose $a^y = x$, so that $y = \log_a x$. According to (6), we have $x = E[yL(a)]$. By (1), this is equivalent to $L(x) = yL(a)$, or

$$\log_a x = \frac{L(x)}{L(a)}. \tag{7}$$

In particular, putting $a = e$, we have

$$\log_e x = L(x), \tag{8}$$

because of the fact that $L(e) = 1$. Also, from (6) we see that

$$e^x = E(x). \tag{9}$$

Formulas (8) and (9) display to some extent the importance of the number $e$, whose definition was given near the end of § 8-2.

We have now completed the framework of the logical development of the properties of logarithms and exponentials as functions. By beginning with the function $L$, as defined in § 8-2 and then introducing its inverse $E$, we have been spared the difficulties of the definition of $a^u$ for irrational $u$ by the method outlined in § 8-1, and we have avoided the difficulties of a direct investigation of the limit indicated in § 8-2, (7). We have used powerful tools, of course; but they are the standard tools of differential and integral calculus, which were already available to us. Some deep questions about these tools do of course remain for investigation in a later course in calculus.

### EXERCISES

1. Prove that $L(a^u) = uL(a)$ for $a > 0$ and $u$ arbitrary. Use the definition of $a^u$.
2. Prove that $[E(u)]^v = E(uv)$ by appropriate use of (5) and (6).
3. Assuming $a > 0$ and the definition (6), prove that $(a^u)^v = a^{uv}$.
4. Prove that $E(u + v) = E(u)E(v)$ by letting $x = E(u)$, $y = E(v)$, $z = E(u + v)$ and using (1) to show that $z = xy$.

## 8-4 Further Discussion of *e*

We have defined $e$ as the unique positive number such that $L(e) = 1$; as a consequence, $E(1) = e$. Since $L'(x) = 1/x$, we see from (7) and (8) in § 8-3 that

$$\frac{d}{dx}\log_a x = \frac{1}{x \log_e a}, \tag{1}$$

and

$$\frac{d}{dx}\log_e x = \frac{1}{x}. \tag{2}$$

Since $\log_e a = (\log_a e)^{-1}$ (see Exercise 5, § 8-1), we can also write

$$\frac{d}{dx}\log_a x = \frac{\log_a e}{x}. \tag{3}$$

Evidently the formula for the derivative of $\log_a x$ is simplest if $a = e$. It is for this reason that mathematicians prefer to use $e$ as a base of logarithms for all theoretical work. Logarithms with base $e$ are called *natural* logarithms, whereas logarithms with base 10 are called *common* logarithms. From now on we follow the standard practice of dropping the basal index $e$ on natural logarithms, so that $\log_e x$ is written $\log x$. We do not drop the basal index in the case of other bases.

There is a widespread usage of the notation $\ln x$ in place of $\log x$ for natural logarithms.

Tables of natural logarithms have been constructed. A small table of this kind is given at the end of this book.

In many calculus texts $e$ is introduced as the limit

$$e = \lim_{t\to 0}(1 + t)^{1/t}. \tag{4}$$

We saw in the early part of § 8-2 that this limit arises in a natural way in connection with the differentiation of logarithms, but we have not yet connected it with the number $e$ as defined in connection with the function $L$. We shall now prove the correctness of (4). Starting with (6) in § 8-3, we write

$$(1 + t)^{1/t} = E\left[\frac{1}{t} L(1 + t)\right]. \tag{5}$$

Since $L(1) = 0$, we can write

$$(1 + t)^{1/t} = E\left[\frac{L(1 + t) - L(1)}{t}\right].$$

Now let $t \to 0$. We see that

$$\frac{L(1 + t) - L(1)}{t} \to L'(1) = 1.$$

Hence, because $E$ is a continuous function, we see that

$$(1 + t)^{1/t} \to E(1) = e.$$

Thus (4) is proved.

It may also be proved that $(1 + t)^{1/t}$ *increases* toward its limit $e$ as $t$ *decreases* toward 0 (see Exercise 2).

EXERCISES

**1.** Use (3) and (6) in § 8-3 to prove that

$$\frac{d}{dx} a^x = a^x \log_e a.$$

Note what this becomes if $a = e$.

**2.** Prove the assertion made just before these exercises, by showing that

$$\frac{d}{dt}(1 + t)^{1/t} < 0$$

when $t > 0$. *Suggestion:* Use (5) to show that the sign of the derivative in question is the same as that of

$$t^{-2}\left[\frac{t}{1 + t} - L(1 + t)\right].$$

Then use the law of the mean (Theorem 2-C) to show that $L(1 + t) > t(1 + t)^{-1}$, thus showing that the derivative in question is negative.

## 8-5 Differentiation Technique

It is desirable for the student to practice differentiation of logarithmic and exponential functions in order to become thoroughly familiar with the formulas and the techniques of using them. The basic formulas are:

$$\frac{d}{dx}\log u = \frac{1}{u}\frac{du}{dx}, \tag{1}$$

$$\frac{d}{dx}e^u = e^u\frac{du}{dx}, \tag{2}$$

where $u$ denotes any differentiable function of $x$. The first of these formulas is obtained from (2) in § 8-4, combined with the chain rule (Theorem 3-E). The second one also is obtained with the aid of the chain rule, for we know from § 8-3 that $E'(x) = E(x) = e^x$.

**Example 1:** Find $y'$ and $y''$ if $y = x^2 \log \cos x$.
Here we use (1) and the rule for products:

$$y' = x^2\frac{1}{\cos x}(-\sin x) + 2x \log \cos x = -x^2 \tan x + 2x \log \cos x.$$

$$y'' = -x^2 \sec^2 x - 2x \tan x + 2x\frac{1}{\cos x}(-\sin x) + 2 \log \cos x$$

$$= -x^2 \sec^2 x - 4x \tan x + 2 \log \cos x.$$

**Example 2:** Use first and second derivatives to study the graph of $y = 3xe^{-x}$.

Here

$$y' = 3xe^{-x}(-1) + 3e^{-x} = 3e^{-x}(1 - x),$$
$$y'' = 3e^{-x}(-1) + 3(1 - x)e^{-x}(-1) = 3e^{-x}(x - 2).$$

Since $e^{-x} > 0$ for all values of $x$, we conclude that the slope is positive if $x < 1$ and negative if $x > 1$. There is a maximum value of $y$ at $x = 1$. From the second derivative we see that the curve is concave upward if $x > 2$, downward if $x < 2$. The graph is shown in Fig. 8-7.

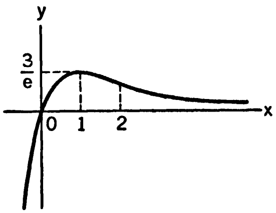



**Fig. 8-7**

It is often useful to know the relative orders of magnitude of the functions $a^x$, $x^n$ ($n$ a positive integer), and $\log_a x$ when $x$ is very large and $a > 1$. All three functions become very large in value, but if $x$ is sufficiently large (just how large depends on $n$ and $a$), we can show that

$$\log_a x < x^n < a^x. \tag{3}$$

Even more is true, namely,

$$\lim_{x \to +\infty} \frac{a^x}{x^n} = +\infty, \tag{4}$$

and

$$\lim_{x \to +\infty} \frac{x^n}{\log_a x} = +\infty. \tag{5}$$

The proofs of (4) and (5) are easily given by methods to be developed later in this book (see l'Hospital's rule, § 14-5).

For reference it is perhaps well to list the formulas

$$\frac{d}{dx} \log_a u = \frac{1}{\log_e a} \frac{1}{u} \frac{du}{dx}, \tag{6}$$

$$\frac{d}{dx} a^u = a^u(\log_e a) \frac{du}{dx}. \tag{7}$$

For the most part, however, we are concerned with the situation when $a = e$.

For some purposes it is convenient to use differentials instead of derivatives. Then we have

$$d \log u = \frac{du}{u}, \qquad de^u = e^u\, du. \tag{8}$$

## EXERCISES

**1.** Find $y'$ in each case.

(a) $y = 5 \log (x^2 + 9)$.

(b) $y = \log (2ax - x^2)$.

(c) $y = \log \sin x$.

(d) $y = \log \tan 3x$.

(e) $y = x^2 \log x$.

(f) $y = x \log (16 - x^2)$.

(g) $y = (\log x^2)^2$.

(h) $y = \dfrac{(\log x)^2}{x}$.

(i) $y = \log \sec 2x$.

(j) $y = \log (\sec 4x + \tan 4x)$.

**2.** Find $y'$ and $y''$ in each case.

(a) $y = x^2 e^{-x}$.

(b) $y = xe^{-x^2}$.

(c) $y = e^x - e^{-x}$.

(d) $y = e^{1/x}$.

(e) $y = e^{\sin x}$.

(f) $y = e^{x \log x}$.

**3.** Find $y'$ in each case.

(a) $y = \log (\log x)$.

(b) $y = \cos (\log x)$.

(c) $y = x^3 \log x - \dfrac{1}{3} x^3$.

(d) $y = x \sin (\log x) - x \cos (\log x)$.

(e) $y = e^{ax}(ax - 1)$.

(f) $y = \log \dfrac{e^x}{1 + e^x}$.

(g) $y = \log \left(\dfrac{e^x - e^{-x}}{e^x + e^{-x}}\right)^{1/2}$.

(h) $y = \tan^{-1} e^{2x}$.

(i) $y = \sin^{-1} (e^{-x/2})$.

(j) $y = x + 2 \log (1 + \sqrt{1 + e^{-x}})$.

**4.** Find $y'$ in each case. Where possible, use properties of logarithms to simplify the expression before differentiating, especially to avoid differentiation of fractional powers and of quotients. See (a) and (b).

(a) $y = \log (1 - x^2)^{1/2} = \dfrac{1}{2} \log (1 - x^2)$.

(b) $y = \log \left(\dfrac{1 + \cos x}{1 - \cos x}\right)^{1/2} = \dfrac{1}{2} \log (1 + \cos x) - \dfrac{1}{2} \log (1 - \cos x)$.

(c) $y = \log \dfrac{x}{1 + x^2}$.

(d) $y = \log \sqrt{\dfrac{1 + x}{1 - x}}$.

(e) $y = \log \sqrt{5 - 2x + 3x^4}$.

(f) $y = \log \dfrac{x + \sqrt{x^2 - a^2}}{a}$.

(g) $y = \log \dfrac{\sqrt{x^2 + a^2} + x}{\sqrt{x^2 + a^2} - x}$.

(h) $y = \log \dfrac{2 \tan x + 1}{\tan x + 2}$.

**5.** (a) If $y = A \sin \log x + B \cos \log x$, prove that $x^2y'' + xy' + y = 0$.

(b) If $y = e^{-x} \sin x$, show that $d^4y/dx^4 + 4y = 0$.

**6.** (a) For what positive $x$ is $y = x^2\, 10^{-x}$ a relative maximum?

(b) Draw the graph and locate the two points of inflection.

7. (a) Show that $y = x^{-1} \log_a x$ has its maximum value at $x = e$ for each choice of $a$, provided $a > 1$.
(b) Sketch the graph and show that the point of inflection is at $x = e^{3/2}$.

8. Graph each curve with the aid of $y'$ and $y''$. Answer the questions as put for each curve.
(a) $y = x^3 e^{-x}$. Locate critical and inflection points.
(b) $y = x \log x$. Find the minimum value of $y$. What happens to $y$ as $x \to 0$?
(c) $y = e^{-x^2}$. Find the points of inflection. Show that, of all rectangles which have two corners on the $x$-axis and two on the curve, that one has the greatest area two of whose corners are at the points of inflection.
(d) $y = x/\log x$. Are there any asymptotes? What is the smallest positive value of $y$? What is the slope at the origin? Is there a point of inflection?
(e) $y = [1 + e^{(x-R)/a}]^{-1}$, $R$ and $a$ positive. Find the point of inflection. Are there asymptotes? This curve has been used in describing the distribu- of charge in an atomic nucleus.

9. If two parallel wires of a transmission line, each of radius $r$, are $h$ units apart, the magnetic flux per unit length between them is proportional to $\log [(h - r)/r]$. Draw a graph showing the way in which the magnetic flux varies as a function of $r$. What is the concavity of the graph when $0 < r < h/2$?

10. Show by mathematical induction that

(a) $\dfrac{d^n}{dx^n} \log x = (-1)^{n-1} \dfrac{(n-1)!}{x^n}$.

(b) $\dfrac{d^n}{dx^n} \log (1 - x) = -\dfrac{(n-1)!}{(1-x)^n}$.

(c) $\dfrac{d^n}{dx^n} (xe^x) = (x + n)e^x$.

(d) $\dfrac{d^n}{dx^n} (xe^{-x}) = (-1)^n(x - n)e^{-x}$.

(e) $\dfrac{d^{n+1}}{dx^{n+1}} (x^n \log x) = \dfrac{n!}{x}$.

11. If $x = \sqrt{a^2 - y^2} - a \log \dfrac{a + \sqrt{a^2 - y^2}}{y}$, show that $y' = y(a^2 - y^2)^{-1/2}$.

12. To differentiate an expression such as $y = x^x$, in which both the base and exponent are variable, it is convenient to begin by taking the logarithm of each side: $\log y = x \log x$. From this equation we may calculate $y'$. Or, alternatively, we may write $x^x = e^{x \log x}$. Find $y'$ and use it to find the minimum value of $x^x$, assuming $x > 0$.

## 8-6 Exponential Growth or Decay

There are many interesting situations in which a variable quantity $y$ changes with time according to the equation

$$\frac{dy}{dt} = ky, \tag{1}$$

where $k$ is a constant, either positive or negative. Such situations occur in chemistry, biology, economics, and in various types of retarded motion.

We get (1) if $y$ depends on $t$ by the formula $y = Ae^{kt}$, where $A$ is some constant. This is clear, by differentiation. Conversely, if we know that $y$ changes in accordance with (1), then

$$\frac{dy}{y} = k\,dt \tag{2}$$

so long as $y \neq 0$. Let us suppose $y > 0$. Then, by antidifferentiation we have

$$\log y = kt + C,$$

where $C$ is some constant. If we suppose that $y = y_0$ when $t = 0$, then $\log y_0 = C$, and so

$$\log y - \log y_0 = \log \frac{y}{y_0} = kt.$$

This is equivalent to

$$\frac{y}{y_0} = e^{kt}, \quad \text{or} \quad y = y_0 e^{kt}.$$

Evidently $y$ decreases with increasing time if $k < 0$, and increases if $k > 0$.

**Example:** The mass of a radioactive body decreases at a rate proportional to the existing mass. If $\frac{1}{4}$ of the initial mass is lost during the first day, how long will it take for $\frac{1}{2}$ the initial mass to be lost in the process of radioactive decay?

Let $m$ be the mass at time $t$, with $m = m_0$ when $t = 0$. Take 1 day as the unit of time. The units for $m$ do not matter, since we deal only with ratios. We are told that

$$\frac{dm}{dt} = km, \quad \text{or} \quad \frac{dm}{m} = k\,dt.$$

Then, just as in the foregoing discussion,

$$\log \frac{m}{m_0} = kt. \tag{3}$$

As yet we do not know the value of $k$. But, since $m = \frac{3}{4}m_0$ when $t = 1$, we see that $\log \frac{3}{4} = k$. Now we can put $m = \frac{1}{2}m_0$ and solve for $t$:

$$t = \frac{\log \frac{1}{2}}{\log \frac{3}{4}} = \frac{\log 2}{\log 4 - \log 3}.$$

From a table of natural logarithms we find

$$t = \frac{0.69315}{0.28768} = 2.409 \cdots.$$

Thus it takes about 2.4 days for half the original mass to be lost.

*Continuous Compounding of Interest*

If a sum of \$P is placed at interest at the nominal interest rate of 6 per cent, compounded semiannually, the accumulated sum, or compound amount, after $t$ years, is

$$S = P(1.03)^{2t}.$$

If the nominal interest rate is $100r$ per cent ($r = 0.06$ for 6 per cent), and if interest is compounded $n$ times a year, the accumulated sum after $t$ years is

$$S = P\left(1 + \frac{r}{n}\right)^{nt}. \tag{4}$$

Let us now suppose that $n$ is increased indefinitely, so that interest is compounded more and more frequently. What happens to $S$ in formula (4) as $n \longrightarrow \infty$? The answer to this question involves the number $e$. Let us write $h_n = r/n$. Then

$$\left(1 + \frac{r}{n}\right)^{nt} = \left[(1 + h_n)^{1/h_n}\right]^{rt}.$$

From (4) in § 8-4 we know that

$$(1 + h_n)^{1/h_n} \longrightarrow e,$$

because $h_n \longrightarrow 0$. Hence, in the limit as $n \longrightarrow \infty$, (4) is replaced by the formula

$$S = Pe^{rt}. \tag{5}$$

Thus we see that, as interest is compounded more and more frequently, the accumulated amount at interest tends to grow exponentially with time. This limiting type of growth is called growth by *continuous compounding* of interest.

*Continuous Dilution of Mixtures*

Suppose a large tank contains $V$ gallons of sea water. In order to reduce the salinity of the water, pure fresh water is run into the tank at the rate of $c$ gallons per minute. The water in the tank is kept thoroughly mixed at all times (this is an idealization of the state of affairs), and the mixture is drawn off at the rate of $c$ gallons per minute, so that the volume of mixture in the tank remains constant. We shall show that the salinity of the water decreases exponentially.

Let $v$ be the number of gallons of mineral salts in the tank at time $t$. In the short interval of time from $t$ to $t + \Delta t$, $c\,\Delta t$ gallons of mixture flow out. Since the proportion of mineral salts in the mixture during this time is approximately $v/V$, the volume of salts carried out is approximately $vc\,\Delta t/V$. Thus, approximately,

$$\Delta v = -\frac{vc\,\Delta t}{V}, \quad \text{or} \quad \frac{\Delta v}{\Delta t} = -\frac{cv}{V}.$$

This approximation becomes better and better as $\Delta t \to 0$, so that the accurate description of the situation is provided by the equation

$$\frac{dv}{dt} = -\frac{cv}{V}.$$

Now the salinity $s$ of the mixture is measured by the ratio $s = v/V$. We see that $v = sV$, and hence

$$V\frac{ds}{dt} = -cs, \quad \text{or} \quad \frac{ds}{dt} = -\frac{c}{V}s. \tag{6}$$

In view of the discussion earlier in this section, this shows that $s$ decreases exponentially.

*Tension in a Rope Around a Rough Cylinder*

It is a matter of common experience that a large force pulling on one end of a rope can be balanced by a small force pulling on the other end, if the rope is snubbed around a rough surface. Here we shall investigate the relation between friction and tension in the case of a rope wound partially or entirely around a rough circular cylinder.

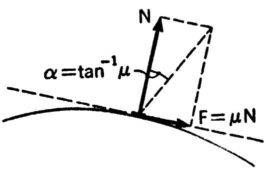


Fig. 8-8

First we recall the basic law of friction. Suppose a small object is in contact with a rough surface, and experiences a force of magnitude $N$ at right angles to the surface as a result of the contact (see Fig. 8-8). This is called the *normal reaction* on the object. If now an attempt is made to move the object along the surface by a force applied tangentially, the friction due to the roughness of the surface will oppose this tangential force up to the maximum amount available from friction. This maximum amount is $F = \mu N$, where $\mu$ is a proportionality factor called the *coefficient of friction;* it is independent of $N$, but depends on the physical characteristics of the surface and the object under consideration. Note that $\mu = \tan \alpha$, where $\alpha$ is the angle marked in Fig. 8-8.

Now consider the situation of a rope around a cylinder, as in Fig. 8-9. We ignore the weight of the rope as a factor in the situation. With a fixed tension $T_0$ where one end of the rope comes off the cylinder, let the tension $T_1$ where the other end comes off be increased until $T_1$ is barely balanced by $T_0$ and the effect of friction. With notation as shown in Fig. 8-9, our basic problem is to discover how the tension $T$ at $P$ depends upon the angular coordinate $\theta$ which locates $P$. In order to get at the problem, consider a small segment of the rope from $\theta$ to $\theta + \Delta\theta$. This segment is in equilibrium under the influence of the tensions $T$ and $T + \Delta T$, at the two ends, and a force $R$ combining the effects of friction and the normal reaction of the cylinder. Let $\beta$ be the angle between the direction of $R$ and the direction $OP$ (see Fig. 8-9). In view of our discussion of friction in connec-

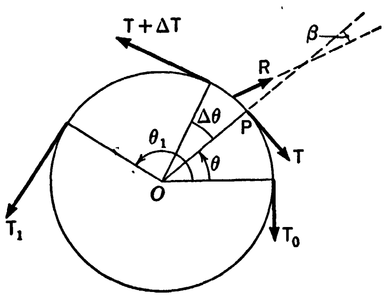



**Fig. 8-9**

tion with Fig. 8-8, it seems plausible to assume that $\tan \beta \to \mu$ as $\Delta\theta \to 0$; here $\mu$ is the coefficient of friction.

The conditions of equilibrium require that the forces along the tangential direction at $P$ shall balance; likewise the forces in the direction of $OP$ must balance. Hence we obtain the two equations

$$T + R \sin \beta = (T + \Delta T) \cos \Delta\theta,$$
$$R \cos \beta = (T + \Delta T) \sin \Delta\theta.$$

We eliminate $R$:

$$\frac{(T + \Delta T) \cos \Delta\theta - T}{(T + \Delta T) \sin \Delta\theta} = \tan \beta.$$

Next we divide numerator and denominator on the left by $\Delta\theta$ and regroup slightly:

$$\frac{T\left(\dfrac{\cos \Delta\theta - 1}{\Delta\theta}\right) + \dfrac{\Delta T}{\Delta\theta} \cos \Delta\theta}{(T + \Delta T)\, \dfrac{\sin \Delta\theta}{\Delta\theta}} = \tan \beta.$$

We are now ready to see what happens when $\Delta\theta \to 0$. From (1) and (2) in § 4-2 we know that

$$\frac{\sin \Delta\theta}{\Delta\theta} \to 1 \quad \text{and} \quad \frac{\cos \Delta\theta - 1}{\Delta\theta} \to 0.$$

Also, $\Delta T \to 0$, $\cos \Delta\theta \to 1$, and $\Delta T/\Delta\theta \to dT/d\theta$. Therefore, since $\tan \beta \to \mu$, we obtain

$$\frac{dT/d\theta}{T} = \mu, \quad \text{or} \quad \frac{dT}{d\theta} = \mu T. \tag{7}$$

From (7) we find at once

$$\frac{dT}{T} = \mu\, d\theta, \qquad \log T = \mu\theta + C.$$

Since $T = T_0$ when $\theta = 0$, we have $C = \log T_0$, and so

$$T = T_0 e^{\mu\theta}. \tag{8}$$

This equation gives the value of $T_1$ by putting $\theta = \theta_1$.

Typical values of $\mu$ range from 0.3 to 0.5 for rope on wood, depending on the particular compositions and textures.

### EXERCISES

1. In a certain chemical reaction, the rate of change of concentration of a substance is proportional to the concentration itself. If the concentration is 1 part in 100 at $t = 0$, and 4 parts in 1000 five minutes later, find the concentration (in parts per 100) as a function of time.

2. In a chemical decomposition, the rate of decomposition of an original 25 kilograms of substance $A$ is proportional to the amount not decomposed. If the mass is reduced to 10 kilograms in 3 hours, when will 24 kilograms be decomposed?

3. If a current from a battery is flowing in a circuit with resistance $R$ and inductance $L$, and the battery is suddenly cut from the circuit, the current $i$ subsequently obeys the law $L\,(di/dt) + Ri = 0$.
(a) Express $i$ as a function of $t$ if $i = i_0$ when $t = 0$.
(b) Taking $R = 1.2$ ohms, $L = 1$ henry, and $i_0 = 5$ amperes, find the number of time units (seconds) until $i = 0.01$ ampere.

4. In 1930 the population of a city was 80,000. In 1950 it was 100,000. If the rate of increase of the population is proportional to the population, (a) what will the population be in 1980? (b) In what year will it be 200,000?

5. The bacteria in a certain culture increase according to the law $dN/dt = kN$. If $N = 3000$ at the outset, and $N = 6000$ when $t = 5$, find (a) $N$ when $t = 1$ and (b) $t$ when $N = 60{,}000$.

6. (a) If a tank holds 5000 gallons of a saline mixture, how many gallons of fresh water must be run into the tank in order to reduce the salinity to 50 per cent of its initial value, following the procedure described in the text?
(b) If fresh water flows in at 50 gallons per minute, by what factor is the salinity reduced in one hour?

7. If $x$, starting from some positive value $x_0$ at $t = 0$, increases or decreases according to the law $dx/dt = kx$, where $k$ is a constant, and if $x_n$ is the value of $x$ when $t = nt_1$ ($t_1$ fixed, $> 0$), show that $x_0, x_1, x_2, \cdots$ is a geometrical progression.

8. Newton's law of cooling states that the difference $x$ between the temperature of a body and that of the surrounding air decreases at a rate proportional to this difference. If $x = 100°$ when $t = 0$ and $x = 40°$ when $t = 40$ minutes, find (a) when $x = 70°$; (b) when $x = 16°$; (c) the value of $x$ when $t = 20$.

9. A flywheel spinning about a shaft is slowed down by friction at a rate proportional to the speed of rotation, so that $dp/dt = -kp$, where $k$ is a positive constant and $p$ is the angular velocity of the flywheel. If the initial angular velocity is 1600 revolutions per minute, and if the velocity is halved in 2 minutes, find (a) the angular velocity after $t$ minutes; (b) the time when $p = 100$ revolutions per minute; the number of radians through which the flywheel has turned in $t$ minutes.

10. If $\mu = 0.5$ for a yacht hawser around a wharf post, how many turns of the rope around the post are necessary in order that a man holding the rope can withstand a pull 100 times as great as that of which he is capable?

11. A 60-pound weight is fastened to one end of a rope. The rope goes straight up, over a horizontal spar of circular cross section, and comes straight down to where a man is standing. If the coefficient of friction between rope and spar is 0.35, how heavy is the man if he can just barely lift himself on the rope without raising the 60-pound weight?

12. In formula (5) $P$ is called the *present value.* Solve the equation for $P$. If the timber on a certain tract will bring $\$100 \cdot e^{\sqrt{t}/2}$ when cut $t$ years from the present, for what value of $t$ is the present value of the timber greatest, assuming that interest is compounded continuously at the nominal rate of 5 per cent per year?

13. If a timber tract costs \$1088 to plant and if the cut timber will bring $\$400 \cdot e^{\sqrt{t}/2}$ after $t$ years, show that the tract will earn the highest nominal rate of interest upon the initial investment if the timber is cut in 16 years. What is this highest rate? Assume $e = 2.72$, and consider that interest is compounded continuously.

14. A man saves at the constant rate of \$1.00 a day, and invests his money. If one thinks of the savings as going into his account continuously, and if interest is earned at the rate of 4 per cent, compounded continuously, how long will it take the man to accumulate \$10,000? Express his savings $x$ in $t$ years as a function of $t$.

15. A piece of real estate worth \$20 billion in 1956 is alleged to have been worth \$20 in 1636. What rate of interest, continuously compounded, would yield this same increase in the same time?

16. A room of volume 12,000 cubic feet had the ventilators closed, and the carbon dioxide content of the air in the room was 0.12 per cent (by volume). The ventilators were then opened, and fresh air, with 0.04 per cent carbon dioxide content, was pumped into the room at a fixed rate.
(a) If in 10 minutes the proportion of carbon dioxide was down to 0.06

per cent, at what rate was fresh air coming in? Assume perfect mixing of the air at all times.

(b) At this same rate, how long will it take to reduce the carbon dioxide content to 0.05 per cent?

(c) What will the percentage be after 20 minutes?

**17.** A large tank has $V$ gallons in it at time $t$. There is a small leak in the bottom, and water escapes at a rate proportional to $V$. Also, water is piped into the tank at the constant rate of $c$ gallons per minute. Let $V = V_0$ when $t = 0$, and let the leakage rate at that instant be $r_0$. Show that $V$ approaches $cV_0/r_0$ as $t \to \infty$, decreasing toward this limit if $r_0 > c$, and increasing toward it if $r_0 < c$. What is the expression for $V$ as a function of $t$?

**18.** A tank containing $V$ gallons of water has, initially, $x_0$ pounds of salt dissolved in the water. A brine containing $\frac{1}{4}$ pound of salt per gallon is run into the tank at a steady rate of $r$ gallons per minute. The solution is kept well stirred, and the stirred mixture runs out at the rate of $r$ gallons per minute. If there are $x$ pounds of salt in the tank $t$ minutes after the process starts, deduce that $\frac{dx}{dt} = r\left(\frac{1}{4} - \frac{x}{V}\right)$, and from this that $x = \frac{V}{4} - \left(\frac{V}{4} - x_0\right)e^{-rt/V}$.

# CHAPTER IX

# HYPERBOLIC FUNCTIONS

## 9-1 Definitions and Properties of Hyperbolic Functions

It was discovered a long time ago that certain special combinations of exponential functions have interesting properties. Consider the two following functions:

$$F_1(x) = \tfrac{1}{2}(e^x - e^{-x}), \qquad F_2(x) = \tfrac{1}{2}(e^x + e^{-x}).$$

Each of these functions is the derivative of the other:

$$F_1'(x) = F_2(x), \qquad F_2'(x) = F_1(x). \tag{1}$$

It is also easily verified that

$$[F_2(x)]^2 - [F_1(x)]^2 = 1. \tag{2}$$

Now the properties (1) and (2) are somewhat like properties of the sine and cosine functions, for if we let $F(x) = \sin x$, $G(x) = \cos x$, then $F'(x) = G(x)$, $G'(x) = -F(x)$, and $[F(x)]^2 + [G(x)]^2 = 1$. There are many other ways in which the functions $F_1$, $F_2$ resemble the sine and cosine. The standard name for $F_1$ is the *hyperbolic sine*, and $F_2$ is called the *hyperbolic cosine*. The abbreviation of "hyperbolic sine of $x$" is sinh $x$, and cosh $x$ stands for "hyperbolic cosine of $x$." Thus

$$\sinh x = \frac{e^x - e^{-x}}{2}, \qquad \cosh x = \frac{e^x + e^{-x}}{2}. \tag{3}$$

By analogy we go on to define the *hyperbolic tangent* function, and other functions, as follows

$$\tanh x = \frac{\sinh x}{\cosh x}, \qquad \operatorname{ctnh} x = \frac{1}{\tanh x},$$

$$\operatorname{sech} x = \frac{1}{\cosh x}, \qquad \operatorname{csch} x = \frac{1}{\sinh x}.$$

Why the adjective *hyperbolic* here? It is because of circumstances like this: The equations $x = a \cosh t$, $y = a \sinh t$ lead to the equation $x^2 - y^2 = a^2$, because of (2). Note that $\cosh t > 0$ for all $t$. We have here a parametric representation of one branch of a rectangular hyperbola. By contrast, the equations $x = a \cos \theta$, $y = a \sin \theta$ furnish a parametric representation of the circle $x^2 + y^2 = a^2$. A further discussion of the parametric representation of the hyperbola is given in § 9-3.

The hyperbolic functions occur so frequently in practice that it is essential to devote this brief chapter to a summary of the principal facts about them. First we consider the graphs of $\sinh x$, $\cosh x$, and $\tanh x$ (see Fig. 9-1). From (3) it appears that $\sinh x$ is an odd function, while $\cosh x$

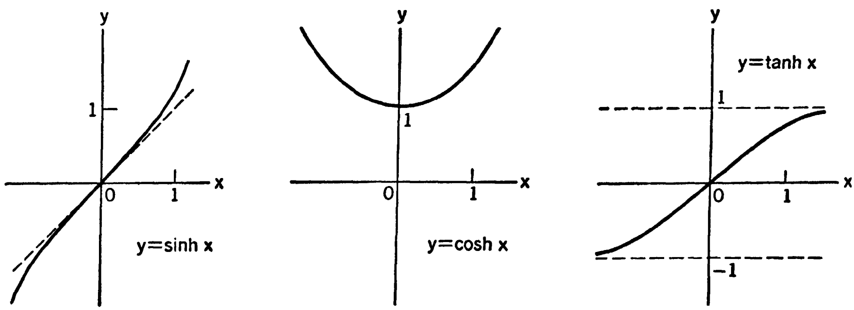



Fig. 9-1

is even. This makes the graph of $y = \sinh x$ symmetric with respect to the origin, while that of $y = \cosh x$ is symmetric with respect to the $y$-axis. For positive $x$, both $\sinh x$ and $\cosh x$ increase as $x$ increases, but $\sinh x$ starts from 0, while $\cosh x$ starts from 1. For very large positive $x$, both functions are large, and very nearly equal, because $e^{-x}$ is then quite small. This causes $\tanh x$ to approach $+1$ as $x \to +\infty$, so that $y = 1$ is an asymptote of the curve $y = \tanh x$. The hyperbolic tangent is an odd function.

*Identities*

There are numerous identities involving hyperbolic functions, each one very much like a corresponding one for trigonometric functions. There are, however, many deviations from trigonometric identities so far as sign is concerned.

One basic identity is (2), which we now write as

$$\cosh^2 x - \sinh^2 x = 1. \tag{4}$$

Certain others may be derived from this. See Exercise 1.

The basic *addition formulas* are:

$$\sinh (x \pm y) = \sinh x \cosh y \pm \cosh x \sinh y. \tag{5}$$

$$\cosh (x \pm y) = \cosh x \cosh y \pm \sinh x \sinh y. \tag{6}$$

An addition formula for the hyperbolic tangent may be derived from (5) and (6). See Exercise 2. Also, half and double variable formulas may be deduced. See Exercise 3.

*Differentiation Formulas*

The following list will be used for reference.

$$\frac{d}{dx} \sinh u = \cosh u \frac{du}{dx}, \tag{7}$$

$$\frac{d}{dx} \cosh u = \sinh u \frac{du}{dx}, \tag{8}$$

$$\frac{d}{dx} \tanh u = \operatorname{sech}^2 u \frac{du}{dx}, \tag{9}$$

$$\frac{d}{dx} \operatorname{ctnh} u = -\operatorname{csch}^2 u \frac{du}{dx}, \tag{10}$$

$$\frac{d}{dx} \operatorname{sech} u = -\operatorname{sech} u \tanh u \frac{du}{dx}, \tag{11}$$

$$\frac{d}{dx} \operatorname{csch} u = -\operatorname{csch} u \operatorname{ctnh} u \frac{du}{dx}. \tag{12}$$

In these formulas $u$ denotes any differentiable function of $x$. We establish (7) and (8) directly from (3), using the chain rule. To get (9), we use the definition of $\tanh u$ as a quotient:

$$\frac{d}{dx}\frac{\sinh u}{\cosh u} = \frac{\cosh u \cdot \cosh u \dfrac{du}{dx} - \sinh u \cdot \sinh u \dfrac{du}{dx}}{\cosh^2 u}$$

$$= \frac{\cosh^2 u - \sinh^2 u}{\cosh^2 u}\frac{du}{dx} = \operatorname{sech}^2 u \frac{du}{dx}.$$

At the last step we used (4) and the definition of $\operatorname{sech} u$.

Further derivations and practice to acquire technique are taken up in the exercises.

## EXERCISES

**1.** Show that (a) $1 - \tanh^2 x = \operatorname{sech}^2 x$, and (b) $\operatorname{ctnh}^2 x - 1 = \operatorname{csch}^2 x$.

**2.** Show that

$$\tanh(x \pm y) = \frac{\tanh x \pm \tanh y}{1 \pm \tanh x \tanh y}.$$

**3.** Prove the following identities:

(a) $\sinh 2x = 2 \sinh x \cosh x$.

(b) $\cosh 2x = \cosh^2 x + \sinh^2 x$.

(c) $\sinh^2 \dfrac{x}{2} = \dfrac{1}{2}(\cosh x - 1)$.

(d) $\cosh^2 \dfrac{x}{2} = \dfrac{1}{2}(\cosh x + 1)$.

**4.** Establish the validity of (10), (11), and (12).

**5.** Verify that the curves $y = \sinh x$, $y = \cosh x$ have positive slope and are concave upward when $x > 0$. What about points of inflection and relative maxima and minima? Calculate the values of $y$ for $x = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}$, using Table II, and prepare your own graphs of these functions on a larger scale than that of Fig. 9-1.

**6.** Discuss slope and concavity of the curve $y = \tanh x$, and use Table II to prepare a graph with 1 inch as unit.

**7.** Show that $(\cosh x \pm \sinh x)^n = \cosh nx \pm \sinh nx$.

**8.** Differentiate each function:

(a) $y = \sinh(3x^2 + 1)$.

(b) $y = \cosh \sqrt{x}$.

(c) $y = \cosh^2 3x$.

(d) $y = \tanh^3 5x$.

(e) $y = \log(\cosh 2x)$.

(f) $y = \tan^{-1}(\sinh x)$.

(g) $y = \dfrac{\sinh x}{1 + \cosh x}$.

(h) $y = \dfrac{2 \cosh 3x}{1 + 2 \sinh 3x}$.

(i) $y = \log(\operatorname{csch} 3x + \operatorname{ctnh} 3x)$.

(j) $y = \log\left(\operatorname{ctnh} \dfrac{x}{2}\right)$.

## 9-2 The Inverse Hyperbolic Functions

We see from the graph of $y = \sinh x$ that for each value of $y$ there is exactly one $x$ such that $\sinh x = y$. This $x$ is denoted by $\sinh^{-1} y$ and called *the inverse hyperbolic sine* of $y$. If we exchange the roles of $x$ and $y$, we have the definition

$$y = \sinh^{-1} x \quad \text{if} \quad x = \sinh y. \tag{1}$$

This inverse function can also be expressed in another way, by use of the logarithm function. If $x = \sinh y$, this is the same as $2x = e^y - e^{-y}$, or

$$e^{2y} - 2xe^y - 1 = 0.$$

This equation can be regarded as a quadratic with $e^y$ as the unknown. Solving, we find

$$e^y = \frac{2x \pm \sqrt{4x^2 + 4}}{2} = x \pm \sqrt{x^2 + 1}.$$

But, since $e^y > 0$, we must take the + sign, and so

$$e^y = x + \sqrt{x^2 + 1}, \qquad y = \log(x + \sqrt{x^2 + 1}).$$

We have thus shown that

$$\sinh^{-1} x = \log(x + \sqrt{x^2 + 1}). \tag{2}$$

To find the derivative of $\sinh^{-1} x$, we start from the second equation in (1) and differentiate:

$$1 = \cosh y \frac{dy}{dx}, \qquad \frac{dy}{dx} = \frac{1}{\cosh y}.$$

But, since $\cosh y > 0$, we see from the basic identity (4) that $\cosh y = \sqrt{1 + x^2}$, and so

$$\frac{d}{dx} \sinh^{-1} x = \frac{1}{\sqrt{1 + x^2}}. \tag{3}$$

The inverse of the hyperbolic tangent can be dealt with in a similar way. See Exercise 1. When we consider the inverse of the hyperbolic cosine, a "principal value" question arises, because for a given $x$ there are two values of $y$ such that $\cosh y = x$, if $x > 1$. We shall take the positive value of $y$ as the principal one, so that $y = \cosh^{-1} x$ means $y > 0$ and $\cosh y = x$. Analogous to (2) we have

$$\cosh^{-1} x = \log(x + \sqrt{x^2 - 1}) \qquad (x \geq 1), \tag{4}$$

and

$$\tanh^{-1} x = \frac{1}{2} \log \frac{1 + x}{1 - x} \qquad (|x| < 1). \tag{5}$$

The differentiation formulas corresponding to (3) are

$$\frac{d}{dx} \cosh^{-1} x = \frac{1}{\sqrt{x^2 - 1}}, \tag{6}$$

$$\frac{d}{dx} \tanh^{-1} x = \frac{1}{1 - x^2}. \tag{7}$$

### EXERCISES

1. (a) Draw the graph of $y = \tanh^{-1} x$, which is the same as that of $x = \tanh y$. Observe that it can be obtained by taking the graph of $y = \tanh x$ on a piece of transparent paper, exchanging the labels on the axes, and looking at the graph from the reverse side of the paper. What asymptotes are there for the graph of $y = \tanh^{-1} x$?
   (b) Derive formulas (5) and (7).

2. Deal with $y = \cosh^{-1} x$ in the manner of Exercise 1, constructing a graph and deriving formulas (4), (6).

3. (a) Construct a graph of $y = \text{ctnh}\, x$, and then a graph of $y = \text{ctnh}^{-1} x$, noting that the latter function is defined when $|x| > 1$.
   (b) Show that

$$\text{ctnh}^{-1} x = \frac{1}{2} \log \frac{x+1}{x-1}, \qquad \frac{d}{dx} \text{ctnh}^{-1} x = \frac{1}{1 - x^2}.$$

4. Show that $\sinh^{-1} \frac{3}{4} = \cosh^{-1} \frac{5}{4} = \tanh^{-1} \frac{3}{5}$. Use Table II to find a two-decimal-place approximate value of these things.

5. Differentiate each function:
   (a) $y = \sinh^{-1} (2x - 1)$.
   (b) $y = \cosh^{-1} (3x + 5)$.
   (c) $y = \tanh^{-1} (2 - 5x)$.
   (d) $y = \tanh^{-1} (\text{sech}\, x)$.
   (e) $y = \cosh^{-1} (\sec x)$.
   (f) $y = \sinh^{-1} (\tan x)$.

## 9-3 Antiderivatives and Integrals

From the differentiation formulas earlier in this chapter we can obtain a number of antiderivative formulas, of which we list the following:

$$\int \sinh u\, du = \cosh u + C. \tag{1}$$

$$\int \cosh u\, du = \sinh u + C. \tag{2}$$

$$\int \text{sech}^2 u\, du = \tanh u + C. \tag{3}$$

$$\int \text{csch}^2 u\, du = -\text{ctnh}\, u + C. \tag{4}$$

$$\int \frac{du}{\sqrt{a^2 + u^2}} = \sinh^{-1} \frac{u}{a} + C. \tag{5}$$

$$\int \frac{du}{\sqrt{u^2 - a^2}} = \cosh^{-1} \frac{u}{a} + C. \tag{6}$$

$$\int \frac{du}{a^2 - u^2} = \frac{1}{a} \tanh^{-1} \frac{u}{a} + C. \tag{7}$$

In (5) and (6) we assume $a > 0$.

We shall now deduce the formula

$$\int \sqrt{a^2 + u^2}\, du = \frac{1}{2}\left(u\sqrt{a^2 + u^2} + a^2 \sinh^{-1} \frac{u}{a}\right) + C. \tag{8}$$

The student should compare the derivation of (8) with the derivation of XI in § 5-5. We let $u = a \sinh t$. Then $du = a \cosh t\, dt$, and

$$\sqrt{a^2 + u^2} = \sqrt{a^2(1 + \sinh^2 t)} = a \cosh t.$$

Therefore $$\int \sqrt{a^2 + u^2}\, du = a^2 \int \cosh^2 t\, dt.$$

Now we use the result of § 9-1, Exercise 3(d), with $t$ in place of $x/2$: $\cosh^2 t = \frac{1}{2}(\cosh 2t + 1)$,

$$\begin{aligned}\int \cosh^2 t\, dt &= \tfrac{1}{4} \int \cosh 2t\, d(2t) + \tfrac{1}{2} \int dt \\ &= \tfrac{1}{4} \sinh 2t + \tfrac{1}{2}t + C.\end{aligned}$$

But by § 9-1, Exercise 3(a),

$$\sinh 2t = 2 \sinh t \cosh t = 2 \cdot \frac{u}{a} \cdot \frac{\sqrt{a^2 + u^2}}{a},$$

and so

$$\int \cosh^2 t\, dt = \frac{1}{2a^2} u\sqrt{a^2 + u^2} + \frac{1}{2} \sinh^{-1} \frac{u}{a} + C.$$

Multiplication by $a^2$ gives us (8).

There is also the formula

$$\int \sqrt{u^2 - a^2}\, du = \frac{1}{2}\left(u\sqrt{u^2 - a^2} - a^2 \cosh^{-1} \frac{u}{a}\right) + C, \tag{9}$$

whose derivation is left for an exercise.

Formulas (8) and (9) enable us to compute certain areas partially bounded by hyperbolas, just as formula XI in § 5-5 enabled us to compute the area within an ellipse.

**Example 1:** Find the area bounded by the right-hand branch of the hyperbola $b^2x^2 - a^2y^2 = a^2b^2$ and the line $x = c$ (which is through the focus, where $c = \sqrt{a^2 + b^2}$).

For review, if necessary, refer to § 3-9. Following the formulation in § 6-5, we see that the required area is

$$A = 2 \int_a^c y\, dx = 2 \frac{b}{a} \int_a^c \sqrt{x^2 - a^2}\, dx.$$

Using (9) and the basic theory of § 6-4, we obtain

$$A = \frac{b}{a}\left[x\sqrt{x^2 - a^2} - a^2 \cosh^{-1} \frac{x}{a}\right]_a^c.$$

Now $\cosh^{-1}(1) = 0$, and $\sqrt{c^2 - a^2} = b$. Hence

$$A = \frac{b^2 c}{a} - ab \cosh^{-1}\left(\frac{c}{a}\right).$$

With $a = 4$, $b = 3$, $c = 5$, for example, this gives $A = 45/4 - 0.7(12) = 2.85$.

The next example shows how we may obtain more antiderivative formulas by substitution. The methods are like those of § 5-4.

**Example 2:** Find $\int \cosh^4 x \sinh x \, dx$ by letting $\cosh x = u$. With this substitution, $du = \sinh x \, dx$, and

$$\int \cosh^4 x \sinh x \, dx = \int u^4 \, du$$
$$= \frac{1}{5} u^5 + C = \frac{1}{5} \cosh^5 x + C.$$

## EXERCISES

**1.** Find each of the indicated antiderivatives.

(a) $\displaystyle\int \sinh^2 2x \cosh 2x \, dx.$ (c) $\displaystyle\int \sinh^3 x \, dx.$

(b) $\displaystyle\int \frac{\sinh x}{\cosh^4 x} \, dx.$ (d) $\displaystyle\int \cosh^5 4x \, dx.$

**2.** Find the indicated antiderivatives by letting $x = a \sinh u$.

(a) $\displaystyle\int \frac{dx}{(x^2 + a^2)^{3/2}}.$ (b) $\displaystyle\int \frac{dx}{x^2\sqrt{x^2 + a^2}}.$

**3.** Find the indicated antiderivatives by letting $x = a \cosh u$.

(a) $\displaystyle\int \frac{dx}{(x^2 - a^2)^{3/2}}.$ (b) $\displaystyle\int \frac{dx}{x^2\sqrt{x^2 - a^2}}.$

**4.** (a) Find $\int \tanh^2 u \, du$ by using Exercise 1(a) from § 9-1 and then (3) of the present section.

(b) Find $\displaystyle\int \frac{x^2 \, dx}{(a^2 + x^2)^{3/2}}$ by letting $x = a \sinh u$ and using the result of (a).

(c) Find $\displaystyle\int \frac{x^2 \, dx}{(x^2 - a^2)^{3/2}}$ by a method analogous to that of (b).

**5.** (a) Derive formula (9).

(b) Show that

$$\int \frac{x^2}{\sqrt{x^2 - a^2}} \, dx = \frac{1}{2}\left(x\sqrt{x^2 - a^2} + a^2 \cosh^{-1} \frac{x}{a}\right) + C$$

and obtain a corresponding formula for the case when $x^2 + a^2$ appears under the radical.

**6.** In Fig. 9-2 the curve $NP$ is $y = b\cosh(x/a)$ and the curve $OQ$ is $y = b\sinh\frac{x}{a}$ ($a$ and $b$ positive).

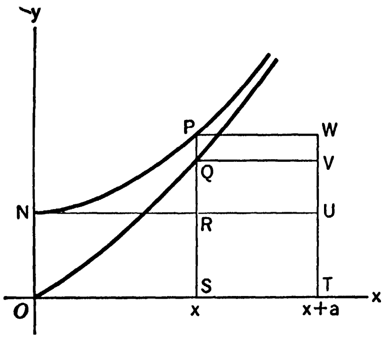



**Fig. 9-2**

(a) If $A_1$ is the area $OSPN$ and $A_2$ is the area $OSQ$, show that $A_1$ = area $QSTV$ and $A_2$ = area $PRUW$.
(b) Show that $A_1/A_2 = \operatorname{ctnh}(x/2a)$ and $A_1 - A_2 = ab(1 - e^{-x/a})$. What happens to the ratio and the difference, respectively, as $x \to +\infty$?

**7.** Consider the branch $x = \sqrt{1 + y^2}$ of the rectangular hyperbola $x^2 - y^2 = 1$, as shown in Fig. 9-3. Let $P$ be the point $(x, y)$. Compute the area $OQPA$, using formula (8). By interpreting part of the answer as the area of the triangle $OPA$, show that the area $OQP$ is $\frac{1}{2}\sinh^{-1} y$, and hence $\sinh^{-1} y$ = area $OP'QP$. Show that this is also equal to $\cosh^{-1} x$. Hence, in the parametric representation $x = \cosh t$, $y = \sinh t$, $t$ can be interpreted as the area $OP'QP$.

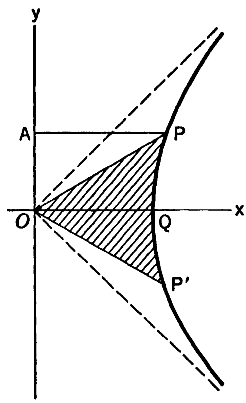



**Fig. 9-3**

**8.** (a) Derive or verify the formula

$$\int \tanh x\, dx = \log(\cosh x) + C.$$

(b) Calculate the area between $y = 1$ and $y = \tanh x$ for $0 \le x \le c$. What does this area approach as $c \to +\infty$?

**9.** (a) Draw the graph of $y = \operatorname{sech} x$. What symmetry and what asymptote do you find?
(b) Derive or verify the formula

$$\int \frac{dx}{\cosh x} = \tan^{-1}(\sinh x) + C.$$

(c) Find the area between $y = \operatorname{sech} x$ and the $x$-axis for $0 \le x \le c$, and find the limit of this area as $c \to +\infty$.

## Review Questions and Problems for Chapters VII, VIII, and IX

### CONCEPTS AND DEFINITIONS

1. What is meant by saying that circle $C_1$ is orthogonal to circle $C_2$? When are two families of circles called orthogonal?

2. State carefully the definition of the radical axis of two circles. Must the circles intersect? What condition must the circles satisfy in order to have a radical axis?

3. What is a coaxal family of circles?

4. When is a family of ellipses confocal? If a family of curves, consisting partly of ellipses and partly of hyperbolas, is confocal, what important relationship exists between the hyperbolas and the ellipses? Can you prove this?

5. What is a homogeneous quadratic form in $x$ and $y$?

6. If one assumes as known the meaning of $a^b$, where $a > 0$ and $b$ is any number, how is $\log_a A$ defined? What is the restriction on $A$, and what is assumed about the behavior of exponentials to make this definition of $\log_a A$ legitimate?

7. How would you attempt to explain exactly what $3^{\sqrt{2}}$ represents, in talking to a high school student?

8. Review in your mind the text's approach, through calculus, to the definitions of exponentials and logarithms.
   (a) What is the definition of $L(x)$?
   (b) What is the definition of $E(x)$? What exactly is the relation between the functions $E$ and $L$?
   (c) How is $a^u$ defined in terms of $E$ and $L$?
   (d) What is the connection between $\log_a x$ and the function $L$?
   (e) What is the definition of the number $e$ in this orbit of ideas?

9. What differential equation is characteristic of exponential growth or decay?

10. What is meant by continuous compounding of interest?

11. Define each of the six hyperbolic functions.

### THEORY

1. Derive the formula for the perpendicular distance from a point to a straight line.

2. Explain how the formula referred to in the foregoing question is used to find equations of the bisectors of the angles formed by two prescribed lines.

3. What is the *normal form* of the equation of a straight line? Explain how it may be used to find the distance between two parallel lines.

4. Work out the equations which express a rotation of axes with rectangular coordinate systems. Use a diagram, and express the coordinates of each system in terms of those in the other.

5. Make a *brief* résumé, without too many details, of the important facts which have been established in the text about quadratic forms and the loci of equations $Ax^2 + 2Bxy + Cy^2 = 1$. Indicate what can be found out without explicitly performing any rotations of axes. Upon what facts, about the coefficients $A$, $B$, $C$ in relation to a rotation of axes, do these findings depend?

6. Outline the main results of a study of the general equation of second degree in $x$ and $y$.

7. Starting from the definition of $L(x)$ by an integral, show that $L(AB) = L(A) + L(B)$ if $A > 0$, $B > 0$. How is it proved that $L(x) \to +\infty$ as $x \to +\infty$ and $L(x) \to -\infty$ as $x \to 0^+$?

8. Justify the fact that there is a unique $x > 0$ for which $L(x) = 1$. What symbol do we regularly use for this value of $x$? How can you obtain a crude estimate for its size?

9. How is it known that for each real $x$ there is a unique positive $y$ such that $L(y) = x$, thus permitting us to define $E$ by $y = E(x)$? Prove that $E(u + v) = E(u)E(v)$, and that $E'(x) = E(x)$.

10. Assuming $a > 0$, define $a^u$ in terms of $E$ and $L$, and prove that $a^{u+v} = a^u a^v$. Now assume $a > 1$ and explain why, for each $x > 0$, there is a unique $y$ such that $a^y = x$, thus defining $y = \log_a x$. Show how to express $\log_a x$ in terms of the function $L$.

11. Use what has been developed to demonstrate that $(1 + t)^{1/t}$ increases toward $e$ as limit as $t$ decreases toward 0.

## PROBLEMS

1. By a translation of axes change the equation $xy - 2x - y - 2 = 0$ to the form $uv =$ constant. What is the curve? Draw it.

2. Simplify the equation $(4x - 3y)^2 = 250x$ by making a rotation of axes so that the line $4x - 3y = 0$ becomes the $u$-axis. Identify the curve and draw it.

3. A line passes through (8,4) and cuts the $y$-axis at $M$, the $x$-axis at $N$. Let $P$ be the mid-point of $MN$. Find the locus of $P$ as the line turns. Identify the curve and find its center of symmetry.

4. (a) Explain why, if one combines the two equations $x^2 - x + 2y = 0$ and $x^2 - 2x + y = 0$ by subtraction to obtain $x + y = 0$, the last equation represents the line through the points of intersection of the curves represented by the first two equations.
(b) Find the straight line through the two points of intersection of the parabolas $x^2 - 4x + \frac{4}{3}y = 0$, $3x^2 - 18x - 4y + 24 = 0$.
(c) Use a method like that of § 7-3 to write an equation of a family of

parabolas, all of which go through the points of intersection of the parabolas in (b). Then select the one which goes through $(0, -1)$.
(d) What happens if you attempt to find a parabola of the family through a point for which $x = 1$ or $x = 4$? Can you account for the result?

**5.** If $B \neq 0$ and $Ax^2 + 2Bxy + Cy^2 = 1$ is an ellipse or a hyperbola, show that the equations of the axes of symmetry will be found by factoring $B(y^2 - x^2) + (A - C)xy$ and setting each factor equal to zero. Use the results developed in § 7-7.

**6.** If $Ax^2 + 2Bxy + Cy^2 = 1$ is an ellipse, show that its area is $\pi(AC - B^2)^{-1/2}$.

**7.** Let $A$ and $B$ be the points $(-4, 0)$, $(4, 0)$, respectively. Let $M$ and $N$ be on the $y$-axis, with $M$ below $N$ and $MN = 4$. Let $P$ be the intersection $AM$ and $BN$. Find the equation of the locus of $P$, and show that the locus is a hyperbola. Find its asymptotes and its axes. (See Problem 5.)

**8.** Consider the equation $Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F = 0$. Suppose it represents a parabola with axis not parallel to the $x$-axis.
(a) Why then is $C \neq 0$?
(b) Show that if we solve for $y$, the solution takes the form $y = px + q \pm \sqrt{rx + s}$, where $p$, $q$, $r$, and $s$ are certain numbers. Hence show that $\dfrac{d^2}{dx^2}[(y'')^{-2/3}] = 0$.

**9.** Given the two families of lines

$$a(2x + y + 3) + b(2x - y + 5) = 0$$
$$h(x - y - 1) + k(3x - 2y - 8) = 0,$$

find the line which belongs to both families without solving to find the common point of either family.

**10.** Find all lines through $(4, -3)$ for which the $x$-intercept is the cube of the $y$-intercept.

**11.** Find all lines through $(2, 7)$ for which the segment cut from the line by axes has length $5\sqrt{2}$.

**12.** Consider the fixed ellipse $b^2x^2 + a^2y^2 = a^2b^2$. If $(x_0, y_0)$ is a point of this ellipse, form the rectangle with sides $x = \pm x_0$, $y = \pm y_0$ and inscribe in it an ellipse with the same axes of symmetry as the original ellipse. Consider $x_0$ as a parameter, $0 < x_0 < a$, and obtain the equation of the family of inscribed ellipses. Draw a number of them and show that they are all tangent to the line $bx + ay = ab$. What are the coordinates of the points of tangency, in terms of the parameter?

**13.** Consider the parabola $2py = x^2$ with directrix $y = -p/2$.
(a) Let $(x_0, y_0)$ be any point on the parabola. With $x_0$ as parameter, write the equation of the family of all tangents to the parabola.
(b) If $(x_1, y_1)$ is another point on the parabola, at which the tangent is perpendicular to the tangent through $(x_0, y_0)$, express $x_1$ in terms of $x_0$.

Then find the intersection of the two tangents. What happens to this point as $x_0$ varies?

**14.** (a) Are there any tangents to the hyperbola $x^2 - 4y^2 = 5$ with slope $\frac{1}{3}$? With slope 1? With slope $-2$?
(b) Consider the family of all straight lines of slope $m$ (where $m$ is fixed). Work out the conditions that a line of this family *shall not* be parallel to an asymptote of the hyperbola in (a) and shall have *just one* point of intersection with the hyperbola. Put the conditions in terms of an inequality which must be satisfied by $m$ and an expression for the square of the $y$-intercept of the line as a function of $m$. As a sample, find the tangents of slope 3 without finding the points of tangency.

**15.** If $e^y + e^x = e^{x+y}$, show that $y' = -e^{y-x}$.

**16.** Draw the curve $y = \log \tan^2 x$, $0 < x < \pi/2$. Find the point of inflection and the slope at that point.

**17.** Draw the curve $y = 3x^2e^{-x^2}$, finding the points where $y$ reaches its largest and smallest values. Obtain the equation from which the points of inflection may be found.

**18.** Find the maximum value of $y$ if $y = 1 - x - e^{-2x}$. Sketch the graph.

**19.** Graph the equation $E = -\dfrac{V}{x \log \dfrac{x}{R}}$, where $V$ and $R$ are positive constants, and $0 < x < R$. What is the minimum value of $E$? This problem occurs in the study of current leakage through the insulation between the conductors of a cylindrical cable.

**20.** Prove that the curve $y = e^{-x}$ is tangent to the curve $y = e^{-x} \sin x$ wherever the two curves have a point in common. Draw $y = e^{-x}$, $y = -e^{-x}$, and $y = e^{-x} \sin x$ on the same coordinate system.

**21.** Assume that the atmospheric pressure at $h$ feet above sea level is $p = 2{,}116e^{-ch}$ pounds per square foot, where $c = (3.8)10^{-5}$. If a plane is 4 miles above the earth and climbing 176 feet per second, what is the rate of change of atmospheric pressure outside the plane?

**22.** If $a > 1$, find the minimum value of $f(x) = x - \log_a x$. For what values of $a$ will the minimum be negative?

**23.** Prove that, for any positive integer $n$, if

$$f_n(x) = e^x - \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}\right),$$

then $f_n(x)$ is positive when $x > 0$, and increases when $x > 0$ and $x$ increases. *Suggestion:* Start with $n = 1$ and use induction. What relation does $f'_{n+1}(x)$ bear to $f_n(x)$?

**24.** Prove that the rule $f'(x) = cx^{c-1}$ is valid if $f(x) = x^c$ and $c$ is irrational. Here we assume $x > 0$. *Suggestion:* Use (6) in § 8-3 to express the meaning of $x^c$. This is the same as $x^c = e^{c \log x}$. Then differentiate, using the known rules for the exponential and logarithm.

**25.** Prove from the definition of $L(x)$ in § 8-2 that

$$\frac{x}{1+x} < L(1+x) < x \quad \text{if} \quad x > 0.$$

Do this by interpreting the integral geometrically and getting upper and lower estimates of its size.

**26.** Show that if $c > 0$, then $\dfrac{\log x}{x^c} \to 0$ when $x \to +\infty$, by reasoning as follows: When $1 < t$, $t^{-1} < t^{b-1}$, where $b = \dfrac{c}{2}$. Then, if $1 < x$,

$$\log x < \int_1^x t^{b-1}\,dt = \frac{x^b - 1}{b} < \frac{x^b}{b}.$$

Explain the first inequality on the line above. Then show how the original assertion follows from what we have.

# CHAPTER X

# THE TECHNIQUE OF INTEGRATION

## 10-1 Indefinite Integrals

In Theorems 6-C and 6-D one finds in precise form the relationship between derivatives and integrals. For practical work with integrals Theorem 6-D is of the utmost importance, because it provides the method by which we calculate the values of the definite integrals which we use in expressing such things as areas, volumes, moments of inertia, work done by forces, and so on. At this point the student should reread § 6-4.

To use Theorem 6-D to calculate a definite integral, we begin by searching for a suitable antiderivative. Up to now we have relied on a comparatively small stock of information about antiderivatives. In order to increase this stock of information and thereby greatly increase our ability to solve a wide variety of problems, we are going to devote this chapter to the systematic development of skill in finding antiderivatives.

Because of the fundamental connection between antiderivatives and integrals, and because of historical usage, antiderivatives are often called *indefinite integrals.* (They are sometimes also called *primitives*, especially in Europe.) The systematic technique of discovering antiderivatives of given functions is called the *technique of integration.* A limited amount of such technique has already been developed in § 5-3, § 5-4, and § 5-5. The student should reread these sections at this time. In particular, Theorem 5-B is the cornerstone of the technique of integration, for practically all such technique, at least in elementary calculus, stems from the use of substitutions.

If $f$ is a function which is continuous on the interval $[a, b]$, Theorem 6-C assures us that the function

$$F(x) = \int_a^x f(t)\,dt$$

is an indefinite integral of $f$, for the theorem states that $F'(x) = f(x)$, which means that $F$ is an antiderivative, or indefinite integral, of $f$. Our present problem, however, is this: Suppose that $f$ is some kind of function whose definition is made in terms of expressions from algebra or trigonometry, or in terms of exponentials and logarithms, or by some finite combination of these types. For the sake of definiteness, even though our terms of reference are not absolutely precise, let us call such functions "elementary." Now if we are presented with some particular elementary function $f$, we would like, if possible, to be able to find an elementary function $F$ which is an indefinite integral of $f$. This is not *always* possible, but it is possible in many cases. The technique of integration proceeds by singling out various classes of elementary functions for which elementary indefinite integrals can be found by suitable devices. We concentrate on the cases of greatest usefulness in this classification.

The student will naturally want to know some examples of *nonelementary* functions. The following functions are not elementary:

$$\int_0^x e^{-t^2}\,dt, \quad \int_0^x \frac{\sin t}{t}\,dt, \quad \int_0^x \frac{dt}{\sqrt{(1-t^2)(4-t^2)}}. \tag{1}$$

That is, there are no functions which are elementary in the sense previously defined, whose derivatives are, respectively,

$$e^{-x^2}, \quad \frac{\sin x}{x}, \quad \frac{1}{\sqrt{(1-x^2)(4-x^2)}}.$$

Yet each of the functions defined by the definite integrals in (1) is useful and interesting. As we progress into more advanced mathematics it becomes more and more necessary to study nonelementary functions.

## 10-2 Commonplace Substitutions

In finding indefinite integrals by substitution the simplest kind of substitution is one which reduces the problem to the form of finding $\int u^n\,du$. We list some standard types.

For $\int (a + bx)^n\,dx$ let $u = a + bx$. (1)

For $\int (a^2 \pm x^2)^n x\,dx$ let $u$ or $u^2$ be $a^2 \pm x^2$. (2)

$$\text{For } \int \sin^n ax \cos ax\, dx \quad \text{let } u = \sin ax. \tag{3}$$

$$\text{For } \int \cos^n ax \sin ax\, dx \quad \text{let } u = \cos ax. \tag{4}$$

In all these cases $n$ need not be an integer. These types were all illustrated in § 5-4, but at that time we could not deal with the case $n = -1$, because the calculus of logarithms had not yet been discussed. Now we know that $d \log u = du/u$, and so

$$\int \frac{du}{u} = \log u + C.$$

**Example 1:** Find $\int \frac{x\, dx}{16 + x^2}$.

We let $u = 16 + x^2$, $du = 2x\, dx$. Then

$$\int \frac{x\, dx}{16 + x^2} = \frac{1}{2} \int \frac{du}{u} = \frac{1}{2} \log (16 + x^2) + C.$$

Since $\log u$ is not defined if $u < 0$, it is well to notice that

$$\frac{d}{du} \log |u| = \frac{1}{u} \quad \text{if} \quad u \neq 0, \tag{5}$$

and hence

$$\int \frac{du}{u} = \log |u| + C \tag{6}$$

is a formula which works for $u < 0$ as well as for $u > 0$. To see that (5) is true when $u < 0$, observe that $|u| = -u$ in that case, and $d \log (-u) = \frac{1}{-u} d(-u) = \frac{du}{u}$.

**Example 2:** Find $\int \tan x\, dx = \int \frac{\sin x}{\cos x}\, dx$.

This is one of our standard types. We let $u = \cos x$, $du = -\sin x\, dx$. Then

$$\int \tan x\, dx = \int \frac{-du}{u} = -\log |\cos x| + C. \tag{7}$$

There is also the formula

$$\int \operatorname{ctn} x\, dx = \log |\sin x| + C, \tag{8}$$

whose derivation we leave to the student.

Types (1)-(4) by no means exhaust the possibilities for commonplace substitutions which reduce a problem to the form $\int u^n\, du$. But it is no use trying to make an elaborate list. Moreover, the scope of simple substitutions is not confined to the form $\int u^n\, du$, but extends to other standard forms, such as

$$\int \frac{du}{\sqrt{a^2 - u^2}} = \sin^{-1}\frac{u}{a} + C,$$

$$\int \frac{du}{a^2 + u^2} = \frac{1}{a}\tan^{-1}\frac{u}{a} + C,$$

$$\int e^u\,du = e^u + C.$$

The exercises provide a varied range for ingenuity and powers of observation. We conclude this section with a few more illustrations. The student may find it convenient to refer to the Table of Integrals in the back of the book. The first 14 integrals in this table are the ones we are starting off with as known at the present stage of development.

**Example 3:** $\int \frac{\log x\,dx}{x} = \int u\,du = \frac{1}{2}u^2 + C$, where $u = \log x$.

**Example 4:** $\int xe^{-x^2}\,dx = -\frac{1}{2}\int e^u\,du = -\frac{1}{2}e^u + C$, where $u = -x^2$.

**Example 5:** $\int \frac{x\,dx}{9 + x^4} = \frac{1}{2}\int \frac{du}{9 + u^2} = \frac{1}{6}\tan^{-1}\frac{u}{3} + C$, where $u = x^2$.

**Example 6:** $\int \frac{\sec^2 x}{\sqrt{4 - \tan^2 x}}\,dx = \int \frac{du}{\sqrt{4 - u^2}} = \sin^{-1}\frac{u}{2} + C$, where $u = \tan x$.

## EXERCISES

**1.** Find each indefinite integral. Check by differentiation.

(a) $\int \tan(3x - 4)\,dx.$

(b) $\int e^{-\sin x}\cos x\,dx.$

(c) $\int \frac{dx}{x\log x}.$

(d) $\int \frac{e^{-x}}{e^{-x} + 2}\,dx.$

(e) $\int \frac{e^x}{1 + e^{2x}}\,dx.$

(f) $\int \frac{\sin 2x\cos 2x}{\sqrt{9 - \cos^4 2x}}\,dx.$

**2.** Proceed as directed in Exercise 1.

(a) $\int \frac{(\log x)^2}{x}\,dx.$

(b) $\int \frac{\cos 2x}{5 - 4\sin 2x}\,dx.$

(c) $\int x\tan x^2\,dx.$

(d) $\int 2x^2\operatorname{ctn} x^3\,dx.$

(e) $\int \frac{e^x + e^{-x}}{e^x - e^{-x}}\,dx.$

(f) $\int \frac{\log(1 + x^2)}{1 + x^2}\,x\,dx.$

## 10-3 Completing the Square. A Reduction Formula

An important device in many integration problems is that of completing the square in a quadratic expression.

**Example 1:** Consider $\int \frac{dx}{\sqrt{6x - 4x^2}}$.

We begin by completing the square in the expression under the radical:

$$6x - 4x^2 = -4(x^2 - \tfrac{3}{2}x + \tfrac{9}{16}) + \tfrac{9}{4}$$
$$= \tfrac{9}{4} - 4(x - \tfrac{3}{4})^2.$$

Thus
$$\int \frac{dx}{\sqrt{6x - 4x^2}} = \int \frac{dx}{\sqrt{\frac{9}{4} - 4(x - \frac{3}{4})^2}}.$$

We may now substitute either $u = x - \frac{3}{4}$ or $u = 2(x - \frac{3}{4})$. With the latter substitution we have $du = 2\,dx$, and our integral becomes

$$\frac{1}{2}\int \frac{du}{\sqrt{\frac{9}{4} - u^2}} = \frac{1}{2}\sin^{-1}\frac{u}{\frac{3}{2}} + C = \frac{1}{2}\sin^{-1}\frac{4x - 3}{3} + C.$$

The device of completing the square is useful in any problem of the types

$$\int \frac{dx}{(ax^2 + bx + c)^n}, \qquad \int \frac{x\,dx}{(ax^2 + bx + c)^n}, \tag{1}$$

and in many other problems where a quadratic expression $ax^2 + bx + c$ with $b \neq 0$ is involved. The exponent $n$ in (1) need not be a positive integer. It might be $\pm\frac{1}{2}$, for instance. The result of completing the square and making a substitution is to give us integrals of the types

$$\int \frac{du}{(Au^2 + B)^n}, \qquad \int \frac{u\,du}{(Au^2 + B)^n} \tag{2}$$

in place of the integrals (1). If we get an integral of the second type in (2) we can substitute $v = Au^2 + B$, $dv = 2Au\,du$; then

$$\int \frac{u\,du}{(Au^2 + B)^n} = \frac{1}{2A}\int \frac{dv}{v^n},$$

and the rest is easy.

For a good technique in completing the square, one may begin by factoring out the coefficient of $x^2$ from the terms in $x^2$ and $x$. For example,

$$2x^2 + 3x + 8 = 2(x^2 + \tfrac{3}{2}x \quad ) + 8.$$

Then one can complete the square inside the parentheses, and make the proper compensation outside. The result in this case is

$$2(x^2 + \tfrac{3}{2}x + \tfrac{9}{16}) + 8 - \tfrac{9}{8} = 2(x + \tfrac{3}{4})^2 + \tfrac{55}{8}.$$

For systematic work we need to know how to evaluate

$$\int \frac{du}{(u^2 + a^2)^n},$$

when $n$ is a positive integer. We already know the result (an inverse tangent) if $n = 1$. We shall now show how to deal with the situation when $n > 1$. We shall prove the formula

$$\int \frac{du}{(u^2 + a^2)^n} = \frac{u}{(2n - 2)(a^2)(u^2 + a^2)^{n-1}} + \frac{2n - 3}{(2n - 2)a^2} \int \frac{du}{(u^2 + a^2)^{n-1}}. \quad (3)$$

This is known as a *reduction formula*, for it shifts our problem from the exponent $n$ to the reduced exponent $n - 1$. The formula is valid if $n \neq 1$.

To prove (3), begin by differentiating $u(u^2 + a^2)^{-n+1}$:

$$d[u(u^2 + a^2)^{-n+1}] = (u^2 + a^2)^{-n+1}\, du - 2(n - 1)(u^2)(u^2 + a^2)^{-n}\, du.$$

Next, write $u^2 = (u^2 + a^2) - a^2$ and put this into the last term:

$$-2(n - 1)(u^2)(u^2 + a^2)^{-n}$$
$$= -2(n - 1)(u^2 + a^2)^{-n+1} + 2(n - 1)(a^2)(u^2 + a^2)^{-n}.$$

Thus

$$d\left[\frac{u}{(u^2 + a^2)^{n-1}}\right] = -\frac{(2n - 3)\, du}{(u^2 + a^2)^{n-1}} + 2(n - 1)(a^2)\frac{du}{(u^2 + a^2)^n}.$$

When we form the indefinite integrals and divide by $(2n - 2)a^2$ we obtain (3).

**Example 2:** Work out $\int \frac{dx}{(x^2 + 9)^3}$.

We have to apply (3) twice. The first use gives us

$$\int \frac{dx}{(x^2 + 9)^3} = \frac{x}{36(x^2 + 9)^2} + \frac{3}{36}\int \frac{dx}{(x^2 + 9)^2}.$$

Also

$$\int \frac{dx}{(x^2 + 9)^2} = \frac{x}{18(x^2 + 9)} + \frac{1}{18}\int \frac{dx}{x^2 + 9}$$
$$= \frac{x}{18(x^2 + 9)} + \frac{1}{54}\tan^{-1}\frac{x}{3} + C.$$

On combining these results we have

$$\int \frac{dx}{(x^2 + 9)^3} = \frac{x}{36(x^2 + 9)^2} + \frac{x}{216(x^2 + 9)} + \frac{1}{648}\tan^{-1}\frac{x}{3} + C.$$

## EXERCISES

Find the indicated indefinite integrals in Exercises 1–10. Check by differentiation.

1. $\int \frac{dx}{\sqrt{9x - 4x^2}}$.
2. $\int \frac{8\, dx}{9x^2 - 12x + 20}$.
3. $\int \frac{x + 1}{\sqrt{8 - 2x - x^2}}\, dx$.
4. $\int \frac{dx}{\sqrt{-x^2 - 5x - 4}}$.
5. $\int \frac{dx}{3x^2 + 14x + 18}$.
6. $\int \frac{x\, dx}{4x^2 - 12x + 13}$.
7. $\int \frac{x - 2}{\sqrt{6x - x^2 - 5}}\, dx$.
8. $\int \frac{2x + 4}{x^2 - 4x + 8}\, dx$.
9. $\int \frac{x\, dx}{\sqrt{4x - x^2}}$.
10. $\int \frac{x\, dx}{9x^2 + 6x + 4}$.

In Exercises 11–16, use the reduction formula (3) after completing the square, or employing other devices, if necessary.

11. $\int \frac{dx}{(4x^2 + 25)^2}$.

12. $\int \frac{dx}{(x^2 + x + 1)^2}$.

13. $\int \frac{x\,dx}{(x^2 - x + 1)^2}$.

14. $\int \frac{x^2\,dx}{(x^2 + 4)^2}$.

15. $\int \frac{dx}{(4x^2 + 16x + 41)^3}$.

16. $\int \frac{x\,dx}{(7x^2 - 14x + 16)^2}$.

## 10-4 Integration of Rational Functions

It is always possible to express the indefinite integral of a rational function in terms of elementary functions. In fact, if $R(x)$ is a rational function, then $\int R(x)\,dx$ can be expressed in terms of rational functions, logarithms of linear and quadratic polynomials, and functions of the form $\tan^{-1}(Ax + B)$, where $A$ and $B$ are constants. We shall justify this assertion by showing how one may proceed in systematic fashion to integrate any rational function.

By definition, a rational function is a quotient of two polynomials. The rational function is called *proper* if the degree of the numerator is less than the degree of the denominator. Otherwise it is *improper*. For example,

$$\frac{x^3}{x^2 + 4} \quad \text{and} \quad \frac{x^2 - x}{1 + x}$$

are improper, while

$$\frac{x}{(x - 1)^2(x + 2)} \quad \text{and} \quad \frac{x^2 + 1}{x(x^2 - 4)}$$

are proper. If we have to integrate an improper fraction, we begin by performing long division until we reach a remainder of degree less than that of the denominator. By this process any improper rational function may be expressed as the sum of a polynomial and a proper rational function.

**Example 1:** Integrate $\int \frac{x^3 - 2x^2}{x^2 + 9}\,dx$.

By long division we find

$$\frac{x^3 - 2x^2}{x^2 + 9} = x - 2 + \frac{-9x + 18}{x^2 + 9}.$$

Thus

$$\int \frac{x^3 - 2x^2}{x^2 + 9}\,dx = \int (x - 2)\,dx - 9\int \frac{x\,dx}{x^2 + 9} + 18\int \frac{dx}{x^2 + 9}.$$

In the second integral we substitute $u = x^2 + 9$, $du = 2x\,dx$. Then

$$\int \frac{x\,dx}{x^2 + 9} = \frac{1}{2}\int \frac{du}{u} = \frac{1}{2}\log u = \frac{1}{2}\log(x^2 + 9).$$

So finally,

$$\int \frac{x^3 - 2x^2}{x^2 + 9}\,dx = \frac{1}{2}x^2 - 2x - \frac{9}{2}\log(x^2 + 9) + 6\tan^{-1}\frac{x}{3} + C.$$

Now let us examine the problem of integrating a proper rational function. We are assuming that all the coefficients are real numbers. There are two types of proper rational functions that can be integrated by methods which we have already developed. The simplest type is

$$\frac{A}{(x - a)^n}, \qquad n = 1, 2, 3, \cdots. \tag{1}$$

A function of this type can be integrated by making the substitution $u = x - a$. Then there is the type

$$\frac{Ax + B}{(x^2 + bx + c)^n}, \qquad n = 1, 2, 3, \cdots \tag{2}$$

in which the quadratic $x^2 + bx + c$ has no real linear factors, i.e., the roots of $x^2 + bx + c = 0$ are imaginary. A function of this type can be integrated by completing the square in the denominator and making an appropriate substitution. This procedure was illustrated in § 10-3. After making the proper substitution, we get integrals of the forms

$$\int \frac{u\,du}{(u^2 + a^2)^n}, \qquad \int \frac{du}{(u^2 + a^2)^n}.$$

For the first form, let $v = u^2 + a^2$. For the second form, use the reduction formula in § 10-3 if $n > 1$; if $n = 1$ we use a standard formula giving us an inverse tangent.

There is a theorem of algebra which guarantees that every proper rational function with real coefficients is expressible in just one way as a sum of functions of the two types (1) and (2). When a proper rational function has been expressed in this way, we say that *it has been decomposed into partial fractions*. We see, therefore, that we can integrate a proper rational function if we can find out how to decompose it into partial fractions.

Here are four samples of decomposition into partial fractions:

$$\frac{7x - 4}{(x - 1)^2(x + 2)} \equiv \frac{2}{x - 1} + \frac{1}{(x - 1)^2} - \frac{2}{x + 2}, \tag{3}$$

$$\frac{x^2 + 6x - 1}{(x - 1)(x - 3)^2} \equiv \frac{3}{2}\cdot\frac{1}{x - 1} - \frac{1}{2}\cdot\frac{1}{x - 3} + \frac{13}{(x - 3)^2}, \tag{4}$$

$$\frac{5}{(x - 1)(x^2 + 4)} \equiv \frac{1}{x - 1} - \frac{x + 1}{x^2 + 4}, \tag{5}$$

$$\frac{-x^3 + 6x^2 + x + 2}{(x^2 - 1)(x^2 + 1)^2} \equiv \frac{1}{x - 1} - \frac{1}{x + 1} - \frac{2}{x^2 + 1} - \frac{x - 2}{(x^2 + 1)^2}. \tag{6}$$

With these examples to illustrate our remarks we shall now explain how to go about expressing a proper rational function in terms of partial fractions.

The first step is to factor the denominator into linear factors and irreducible quadratic factors. A quadratic factor is called irreducible if it cannot be factored into real linear factors. For $ax^2 + bx + c$ this means that $b^2 - 4ac < 0$. Some factors may be repeated. For instance, $x - 1$ is repeated in (3), $x - 3$ is repeated in (4), and $x^2 + 1$ is repeated in (6). The numerator of the fraction need not be factored, but we must make sure that the degree of the numerator is less than that of the denominator.

Each distinct factor of the denominator generates a certain number of terms in the partial fractions decomposition of the given function. If a linear factor $x - a$ is repeated $n$ times, it gives rise in the decomposition into partial fractions, to a sum of terms

$$\frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_n}{(x-a)^n},$$

with constant coefficients $A_1, \cdots, A_n$, some of which may be zero. This is illustrated in the case of (3), where the sum

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

owes its presence to the repeated factor $(x - 1)^2$ in the denominator on the left, while the single term

$$-\frac{2}{x+2}$$

owes its presence to the nonrepeated factor $x + 2$. Similarly, if an irreducible quadratic factor is repeated $m$ times, it gives rise in the decomposition into partial fractions, to a sum of terms

$$\frac{B_1x + C_1}{x^2 + ax + b} + \frac{B_2x + C_2}{(x^2 + ax + b)^2} + \cdots + \frac{B_mx + C_m}{(x^2 + ax + b)^m}.$$

This is illustrated in (5) and (6), in the case of the nonrepeated factor $x^2 + 4$ in (5) and the repeated factor $(x^2 + 1)^2$ in (6). To illustrate the procedure still further, a proper rational fraction with denominator $x^3(x + 5)(x^2 - x + 1)^2$ admits a decomposition into partial fractions in the form

$$\frac{A}{x} + \frac{B}{x^2} + \frac{C}{x^3} + \frac{D}{x+5} + \frac{Ex + F}{x^2 - x + 1} + \frac{Gx + H}{(x^2 - x + 1)^2}.$$

After it has been determined what kind of partial fraction terms may be present in a particular case, there remains the problem of finding the coefficients. This is an algebraic problem, and involves nothing more difficult than the solution of a system of simultaneous linear equations. We illustrate the procedure by examples.

**Example 2:** Consider $\dfrac{2x}{(x-1)^2(x+2)}$.

The decomposition into partial fractions has the form

$$\frac{2x}{(x-1)^2(x+2)} \equiv \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{x+2}. \tag{7}$$

The coefficients $A$, $B$, $C$ are constants; to determine them, the above identity is cleared of fractions.

$$2x \equiv A(x-1)(x+2) + B(x+2) + C(x-1)^2, \tag{8}$$

or, $$2x \equiv (A+C)x^2 + (A+B-2C)x + (-2A+2B+C).$$

We next equate coefficients of like powers of $x$ on the two sides of the identity:

$$\begin{aligned} A \phantom{{}+2B} + C &= 0 \\ A + B - 2C &= 2 \\ -2A + 2B + C &= 0. \end{aligned}$$

The solution of these three equations is found to be $A = \frac{4}{9}$, $B = \frac{2}{3}$, $C = -\frac{4}{9}$. Hence, from (7), we have

$$\frac{2x}{(x-1)^2(x+2)} \equiv \frac{4}{9} \cdot \frac{1}{x-1} + \frac{2}{3} \cdot \frac{1}{(x-1)^2} - \frac{4}{9} \cdot \frac{1}{x+2}.$$

An alternative method of finding the coefficients is often useful. It consists in assigning particular values to $x$ in such a way as to give equations involving just one of the unknown coefficients. A value of $x$ which reduces a linear factor of the denominator to zero is always convenient. In the foregoing example let us set $x = 1$ and $x = -2$ successively. We obtain from the identity (8):

$$2 = 3B \quad \text{and} \quad -4 = 9C.$$

Thus we have at once $B = \frac{2}{3}$, $C = -\frac{4}{9}$. There is no particularly simple value which we can assign to $x$ as an aid in finding $A$. Knowing $B$ and $C$, however, we can find $A$ from any one of the equations obtained by equating the coefficients of like powers of $x$ on the two sides of the identity.

**Example 3:** Consider the fraction

$$\frac{4x^2}{(x-1)^2(x^2-x+1)}.$$

We write

$$\frac{4x^2}{(x-1)^2(x^2-x+1)} \equiv \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{Cx+D}{x^2-x+1},$$

$$4x^2 \equiv A(x-1)(x^2-x+1) + B(x^2-x+1) + (Cx+D)(x-1)^2. \tag{9}$$

Setting $x = 1$ in (9) we see that $4 = B$. We need three equations to find $A$, $C$, and $D$. We can get two such equations by equating the coefficients of $x^3$ and $x^2$, respectively, on the two sides of (9). A third equation could be obtained by considering the coefficients of $x$, but it is more convenient to set $x = 0$ in the identity. We leave it for the student to verify that we get the three equations

$$\begin{aligned} A \quad\; + C \qquad\;\; &= 0 \\ -2A + B - 2C + D &= 4 \\ -\,A + B \qquad\; + D &= 0. \end{aligned}$$

We already know that $B = 4$. It is a simple matter to solve and find $A = 4$, $C = -4$, $D = 0$. Thus the decomposition into partial fractions is completed.

**Example 4:** Derive the integration formula

$$\int \frac{du}{a^2 - u^2} = \frac{1}{2a} \log \left|\frac{a + u}{a - u}\right| + C. \tag{10}$$

This important formula will be used for reference purposes. To prove it, decompose the integrand into partial fractions:

$$\frac{1}{a^2 - u^2} \equiv \frac{A}{a + u} + \frac{B}{a - u},$$

$$1 \equiv A(a - u) + B(a + u).$$

To find $A$ and $B$, first set $u = a$ and then $u = -a$, thus obtaining

$$1 = 2aB, \qquad 1 = 2aA, \qquad \text{or} \qquad A = B = \frac{1}{2a}.$$

Hence
$$\frac{1}{a^2 - u^2} \equiv \frac{1}{2a}\left(\frac{1}{a + u} + \frac{1}{a - u}\right);$$

then
$$\begin{aligned} \int \frac{du}{a^2 - u^2} &= \frac{1}{2a} (\log |a + u| - \log |a - u|) + C \\ &= \frac{1}{2a} \log \left|\frac{a + u}{a - u}\right| + C. \end{aligned}$$

## EXERCISES

Find the indicated indefinite integrals.

1. $\int \frac{4 + x^2}{8 + x}\, dx.$
2. $\int \frac{x^2 + 20}{x^2 - 16}\, dx.$
3. $\int \frac{2 + x^3}{x^2 + 4x}\, dx.$
4. $\int \frac{2x^3 - x + 1}{x(x^2 - 4)}\, dx.$
5. $\int \frac{x^4 + 4x^2 - 4}{(x - 1)(x^2 + 4)}\, dx.$
6. $\int \frac{x^4 + 4x - 16}{(2 - x)^2(4 + x^2)}\, dx.$
7. $\int \frac{x^2}{(x - 2)^3}\, dx.$
8. $\int \frac{x^2 - 6}{x(x - 1)^2}\, dx.$

9. $\int \frac{x^2 + 1}{x(2x - 5)} dx.$

10. $\int \frac{x^2 + 1}{x^3 - 4x} dx.$

11. $\int \frac{dx}{x^2 + x - 30}.$

12. $\int \frac{dx}{4x^2 - 12x + 5}.$

13. $\int \frac{x + 2}{(x^2 - 1)(x^2 + 1)^2} dx.$

14. $\int \frac{x + 1}{x(x^3 - 1)} dx.$

15. $\int \frac{dx}{x^3 + 2x - 3}.$

16. $\int \frac{x^2}{x^4 - 16} dx.$

In the following exercises find the indicated area.

17. Between the curve $y(9 - 4x^2) = 18$ and the $x$-axis, from $x = -1$ to $x = \sqrt{2}$.

18. Between the curve $x(16y^2 - 36) = 216$ and the $y$-axis, from $y = 3$ to $y = 6$.

## 10-5 Integration by Parts

If $u$ and $v$ denote differentiable functions of $x$, we know that

$$d(uv) = u\,dv + v\,du, \quad \text{or} \quad u\,dv = d(uv) - v\,du.$$

Hence
$$\int u\,dv = uv - \int v\,du. \tag{1}$$

This is used as a method of finding $\int u\,dv$ if $\int v\,du$ is easier to find than the first integral. The method is called *integration by parts.*

**Example 1:** Consider $\int x \sin x\,dx$.

If we let $u = x$, $dv = \sin x\,dx$, then $du = dx$, we can take $v = -\cos x$, and then (1) gives

$$\begin{aligned}\int x \sin x\,dx &= -x \cos x + \int \cos x\,dx \\ &= -x \cos x + \sin x + C.\end{aligned}$$

Integration by parts is effective in working out the following types of indefinite integrals:

$$n > 0: \quad \int x^n \sin ax\,dx, \quad \int x^n \cos ax\,dx, \quad \int x^n e^{ax}\,dx.$$

$$n \geq 0: \quad \int x^n \sin^{-1} x\,dx, \quad \int x^n \cos^{-1} x\,dx, \quad \int x^n \tan^{-1} x\,dx.$$

$$n > 0: \quad \int x^m (\log x)^n\,dx, \quad m \neq -1.$$

Here $n$ denotes an integer in each case. Indications of how these types are to be treated are found in the Exercises.

The following illustrative example will show how we may work out $\int e^{ax} \sin bx\,dx$ and the corresponding integral involving $\cos bx$.

**Example 2:** Consider $\int e^{ax} \sin bx\, dx$, letting $u = e^{ax}$, $dv = \sin bx\, dx$. Then $du = ae^{ax}\, dx$; $v$ can be taken to be $-\frac{1}{b}\cos bx$. Then

$$\int e^{ax} \sin bx\, dx = -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b}\int e^{ax} \cos bx\, dx.$$

Right here comes the interesting thing in this problem: Although the new integral is no easier than the old, we can make progress by applying the method again on the new integral. We let $u = e^{ax}$, $dv = \cos bx\, dx$, and get $du = ae^{ax}\, dx$, $v = \frac{1}{b}\sin bx$. Then

$$\int e^{ax} \cos bx\, dx = \frac{1}{b} e^{ax} \sin bx - \frac{a}{b}\int e^{ax} \sin bx\, dx.$$

In spite of appearances, we are *not* going in a circle! We substitute our second result in the earlier equation:

$$\int e^{ax} \sin bx\, dx = -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b}\left(\frac{1}{b} e^{ax} \sin bx - \frac{a}{b}\int e^{ax} \sin bx\, dx\right).$$

Now collect the two terms involving the unknown integral:

$$\left(1 + \frac{a^2}{b^2}\right)\int e^{ax} \sin bx\, dx = -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b^2} e^{ax} \sin bx.$$

Thus $$\int e^{ax} \sin bx\, dx = \frac{e^{ax}}{a^2 + b^2}(a \sin bx - b \cos bx).$$

This gives us *one* indefinite integral of $e^{ax} \sin bx$. To get the complete answer add $C$ on the right.

There are other uses of integration by parts. We shall meet one of them in deriving Taylor's formula with integral remainder, in § 15-3.

## EXERCISES

**1.** Find the following integrals, letting $u$ be the power of $x$ under the integral sign, and taking $dv$ to be the rest of the expression.

(a) $\int xe^x\, dx$.

(b) $\int x^2e^{-x}\, dx$.

(c) $\int x^3e^{2x}\, dx$.

(d) $\int x \cos 2x\, dx$.

(e) $\int x^2 \sin x\, dx$.

(f) $\int x^3 \cos x\, dx$.

**2.** (a) Derive the formula

$$\int x^n e^{ax}\, dx = \frac{1}{a} x^n e^{ax} - \frac{n}{a}\int x^{n-1} e^{ax}\, dx.$$

(b) By repeated use of the formula in (a), find $\int x^4 e^{3x}\, dx$.

(c) If $P(x)$ is a polynomial of degree $n$, explain why $\int P(x)e^{ax}\,dx = Q(x)e^{ax} + C$, where $Q(x)$ is some polynomial of degree $n$.

**3.** Derive the formulas

$$\int x^n \sin ax\,dx = -\frac{1}{a}x^n \cos ax + \frac{n}{a}\int x^{n-1}\cos ax\,dx,$$

$$\int x^n \cos ax\,dx = \frac{1}{a}x^n \sin ax - \frac{n}{a}\int x^{n-1}\sin ax\,dx.$$

**4.** Find the following integrals, taking the power of $\log x$ to be $u$ in each case, with $dv$ equal to the rest of the expression.

(a) $\int \log x\,dx.$

(b) $\int x\log x\,dx.$

(c) $\int (\log x)^2\,dx.$

(d) $\int x^{1/2}\log x\,dx.$

(e) $\int x^{3/2}(\log x)^2\,dx.$

(f) $\int x(\log x)^3\,dx.$

**5.** (a) Derive the formula

$$\int x^m(\log x)^n\,dx = \frac{x^{m+1}(\log x)^n}{m+1} - \frac{n}{m+1}\int x^m(\log x)^{n-1}\,dx,$$

where $m \neq -1$. What is the situation if $m = -1$?

(b) By repeated use of the formula in (a), find $\int x^{10}(\log x)^4\,dx$.

**6.** Find the following integrals, letting $u$ be the inverse trigonometric function in each case, with $dv$ equal to the rest of the expression.

(a) $\int \sin^{-1} x\,dx.$

(b) $\int \tan^{-1} x\,dx.$

(c) $\int \text{ctn}^{-1} x\,dx.$

(d) $\int x\tan^{-1} x\,dx.$

(e) $\int x^2\tan^{-1}\frac{x}{2}\,dx.$

(f) $\int \frac{\tan^{-1} x}{x^2}\,dx.$

**7.** Derive the formulas

$$\int x^n \sin^{-1} x\,dx = \frac{x^{n+1}}{n+1}\sin^{-1} x - \frac{1}{n+1}\int \frac{x^{n+1}}{\sqrt{1-x^2}}\,dx,$$

$$\int x^n \tan^{-1} x\,dx = \frac{x^{n+1}}{n+1}\tan^{-1} x - \frac{1}{n+1}\int \frac{x^{n+1}}{1+x^2}\,dx.$$

**8.** Find the following integrals.

(a) $\int e^{ax}\cos bx\,dx.$

(b) $\int x\sec^2 x\,dx.$

(c) $\int x\csc^2 x\,dx.$

(d) $\int \sin(\log x)\,dx.$

(e) $\int \cos(\log x)\,dx.$

**9.** Derive the formula

$$\int \log\,(x + \sqrt{x^2 + a^2})\,dx = x \log\,(x + \sqrt{x^2 + a^2}) - \sqrt{x^2 + a^2} + C.$$

**10.** Make the substitution $u = \sin^{-1} x$ and find the integral $\int (\sin^{-1} x)^2\,dx$.

## 10-6 Certain Trigonometric Integrals

We begin by deriving formula 15 in the Table of Integrals in the back of the book. To deduce 15 write

$$\int \sec x\,dx = \int \frac{dx}{\cos x} = \int \frac{\cos x\,dx}{\cos^2 x} = \int \frac{\cos x\,dx}{1 - \sin^2 x}.$$

Now let $u = \sin x$, $du = \cos x\,dx$, and use formula (10) of § 10-4:

$$\int \frac{\cos x\,dx}{1 - \sin^2 x} = \int \frac{du}{1 - u^2} = \frac{1}{2}\log\frac{1+u}{1-u} = \frac{1}{2}\log\frac{1+\sin x}{1-\sin x} + C. \quad (1)$$

Next, observe that

$$\frac{1+\sin x}{1-\sin x} = \frac{(1+\sin x)^2}{1-\sin^2 x} = \left(\frac{1+\sin x}{\cos x}\right)^2 = (\sec x + \tan x)^2.$$

The first formula in 15 now follows from (1). We leave the derivation of the second form of 15 as an exercise in trigonometry. Formula 16 may be derived in much the same manner as was 15.

Next we consider what to do about $\int \sin^m x \cos^n x\,dx$, where $m$ and $n$ are integers. The most appropriate procedure depends greatly on the exponents.

If either $m$ or $n$ is an odd positive integer, things work out rather simply. For example, if $n$ is odd and positive, we think of $\cos x\,dx$ as $d(\sin x)$, and let $u = \sin x$.

**Example 1:** In $\int \sin^4 x \cos^3 x\,dx$, let $u = \sin x$, $du = \cos x\,dx$. Then $\cos^2 x = 1 - u^2$, and the result is

$$\int \sin^4 x \cos^3 x\,dx = \int u^4(1 - u^2)\,du = \frac{1}{5}u^5 - \frac{1}{7}u^7 + C$$
$$= \frac{1}{5}\sin^5 x - \frac{1}{7}\sin^7 x + C.$$

To integrate $\sin^2 x$ or $\cos^2 x$ we can use the formulas

$$\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta), \qquad \cos^2\theta = \frac{1}{2}(1 + \cos 2\theta).$$

By repeated use of this method we can integrate higher *even* powers of $\sin x$ or $\cos x$. What is involved is essentially an exercise in trigonometry.

**Example 2:** To deal with $\int \cos^4 x\,dx$ we can write

$$\cos^4 x = \left(\frac{1+\cos 2x}{2}\right)^2 = \frac{1}{4}(1 + 2\cos 2x + \cos^2 2x),$$

$$\cos^2 2x = \frac{1}{2}(1 + \cos 4x),$$

$$\int \cos^4 x\,dx = \int \left(\frac{1}{4} + \frac{1}{2}\cos 2x + \frac{1}{8} + \frac{1}{8}\cos 4x\right) dx$$

$$= \frac{3}{8}x + \frac{1}{4}\sin 2x + \frac{1}{32}\sin 4x + C.$$

This procedure is not very convenient, especially if the power is high. It is often more convenient to use one of the reduction formulas

$$\int \sin^n x\,dx = -\frac{\sin^{n-1} x \cos x}{n} + \frac{n-1}{n}\int \sin^{n-2} x\,dx, \tag{2}$$

$$\int \cos^n x\,dx = \frac{\cos^{n-1} x \sin x}{n} + \frac{n-1}{n}\int \cos^{n-2} x\,dx. \tag{3}$$

The first of these formulas may be derived by taking $u = \sin^{n-1} x$, $dv = \sin x\,dx$, and integrating by parts. We have

$$du = (n-1)\sin^{n-2} x \cos x\,dx, \qquad v = -\cos x,$$

whence

$$\int \sin^n x\,dx = -\sin^{n-1} x \cos x + (n-1)\int \sin^{n-2} x \cos^2 x\,dx.$$

But $\cos^2 x = 1 - \sin^2 x$, and so

$$\int \sin^{n-2} x \cos^2 x\,dx = \int \sin^{n-2} x\,dx - \int \sin^n x\,dx.$$

Thus, $$n\int \sin^n x\,dx = -\sin^{n-1} x \cos x + (n-1)\int \sin^{n-2} x\,dx,$$

and from this result (2) follows at once. The derivation of (3) is entirely similar.

**Example 3:** Find $\int \cos^4 x\,dx$, using (3).
A first application of (3) yields the result

$$\int \cos^4 x\,dx = \frac{\cos^3 x \sin x}{4} + \frac{3}{4}\int \cos^2 x\,dx.$$

Now apply (3) again, and note that the zero power of $\cos x$ is unity:

$$\int \cos^2 x\,dx = \frac{\cos x \sin x}{2} + \frac{1}{2}\int dx.$$

Therefore

$$\int \cos^4 x\,dx = \frac{\cos^3 x \sin x}{4} + \frac{3}{4}\left(\frac{\cos x \sin x}{2} + \frac{1}{2}x\right) + C.$$

Powers of the other trigonometric functions may also be dealt with by reduction formulas. We refer the student to formulas 24-30 in the Table of Integrals at the end of the book. While these reduction formulas afford a direct and systematic procedure for integrating powers of trigonometric functions, it sometimes happens that a substitution will accomplish the same result readily. For instance, powers of tan $x$, or positive even powers of sec $x$, may be integrated by using the substitution $u = \tan x$ and noting that $du = \sec^2 x\, dx = (1 + u^2)\, dx$. Similarly, the substitution $u = \operatorname{ctn} x$ may be used to integrate powers of ctn $x$ or positive even powers of csc $x$. For examples of the uses of these substitutions, see Exercises 5 and 6.

Various trigonometric identities are often useful in dealing with trigonometric integrals. For instance, to integrate $\tan^2 x \cos x$, observe that

$$\tan^2 x \cos x = \frac{\sin^2 x}{\cos x} = \frac{1 - \cos^2 x}{\cos x} = \sec x - \cos x.$$

Therefore

$$\int \tan^2 x \cos x\, dx = \int \sec x\, dx - \int \cos x\, dx$$

$$= \log|\sec x + \tan x| - \sin x + C.$$

Certain definite integrals of powers of sin $x$ and cos $x$ occur quite often, and it is convenient to have formulas for their values. We distinguish between even and odd powers.

If $n = 2, 4, 6, \cdots,$

$$\int_0^{\pi/2} \sin^n x\, dx = \int_0^{\pi/2} \cos^n x\, dx = \frac{1\cdot 3\cdot 5 \cdots (n-1)}{2\cdot 4\cdot 6 \cdots n}\frac{\pi}{2}; \tag{4}$$

if $n = 3, 5, 7, \cdots,$

$$\int_0^{\pi/2} \sin^n x\, dx = \int^{\pi/2} \cos^n x\, dx = \frac{2\cdot 4\cdot 6 \cdots (n-1)}{1\cdot 3\cdot 5 \cdots n}. \tag{5}$$

The formulas are derived from the reduction formulas (2) and (3). If $n$ is a positive integer greater than unity, and we integrate between limits 0 and $\pi/2$ in formulas (2) and (3), the integrated terms disappear at both limits, and therefore

$$\int_0^{\pi/2} \sin^n x\, dx = \frac{n-1}{n}\int_0^{\pi/2} \sin^{n-2} x\, dx,$$

$$\int_0^{\pi/2} \cos^n x\, dx = \frac{n-1}{n}\int_0^{\pi/2} \cos^{n-2} x\, dx.$$

By repeated use of these formulas, we can reduce the exponents until we arrive at one of the integrals

$$\int_0^{\pi/2} \sin x\, dx, \qquad \int_0^{\pi/2} \cos x\, dx, \qquad \int_0^{\pi/2} dx,$$

whose values are respectively 1, 1, $\pi/2$. If, for instance, $n$ is even,

$$\begin{aligned}\int_0^{\pi/2} \sin^n x\,dx &= \frac{n-1}{n}\int_0^{\pi/2} \sin^{n-2} x\,dx \\ &= \frac{(n-1)(n-3)}{n(n-2)}\int_0^{\pi/2} \sin^{n-4} x\,dx \\ &= \cdots = \frac{(n-1)(n-3)\cdots 3\cdot 1}{n(n-2)\cdots 4\cdot 2}\int_0^{\pi/2} dx,\end{aligned}$$

and so we have the first result in (4). The other cases of (4) and (5) should be worked out in a similar fashion by the student.

Trigonometric integrals of the types

$$\int \sin mx \cos nx\,dx,$$

$$\int \sin mx \sin nx\,dx,$$

$$\int \cos mx \cos nx\,dx,$$

$$m \neq n,$$

may be handled easily with the aid of the following formulas:

$$\begin{aligned}2 \sin A \cos B &= \sin (A + B) + \sin (A - B), \\ 2 \sin A \sin B &= \cos (A - B) - \cos (A + B), \\ 2 \cos A \cos B &= \cos (A - B) + \cos (A + B).\end{aligned}$$

**Example 4:** Find $\int \sin 2x \cos 3x\,dx$.

We set $A = 2x$, $B = 3x$. Then

$$\begin{aligned}\int \sin 2x \cos 3x\,dx &= \frac{1}{2}\int (\sin 5x - \sin x)\,dx \\ &= -\frac{1}{10}\cos 5x + \frac{1}{2}\cos x + C.\end{aligned}$$

## EXERCISES

1. Derive the two forms of formula 16 in the Table of Integrals.
2. Work out each of the following indefinite integrals by using an appropriate substitution.

(a) $\int \cos^2 x \sin^5 x\,dx$.

(b) $\int \cos^3 5x\,dx$.

(c) $\int \sec^4 x \sin^3 x\,dx$.

(d) $\int \cos^3 \frac{x}{2} \csc^4 \frac{x}{2}\,dx$.

(e) $\int \sqrt{\cos x}\,\sin^3 x\,dx$.

(f) $\int \cos^5 (2x - 1) \sin^2 (2x - 1)\,dx$.

**3.** Show that the answers to Examples 2 and 3, though different in appearance, are in fact in agreement.

**4.** Work out $\int \sin^4 x\, dx$ in two different ways, and show that the results are in agreement.

**5.** Work out the indicated indefinite integrals by using the substitution $u = \tan x$. Recall that $\sec^2 x = 1 + \tan^2 x$. This method works well for $\int \tan^n x\, dx$, where $n$ is any integer, and for $\int \sec^m x\, dx$, where $m$ is an even positive integer.

(a) $\int \tan^2 x\, dx.$ (d) $\int \sec^4 x\, dx.$

(b) $\int \tan^3 x\, dx.$ (e) $\int \sec^6 x\, dx.$

(c) $\int \tan^4 x\, dx.$ (f) $\int \sec^4 x \tan^2 x\, dx.$

**6.** The following indefinite integrals can be worked out by a procedure analogous to that of Exercise 5.

(a) $\int \operatorname{ctn}^2 x\, dx.$ (c) $\int \csc^6 x\, dx.$

(b) $\int \operatorname{ctn}^5 x\, dx.$ (d) $\int \operatorname{ctn} 2x \csc^4 2x\, dx.$

**7.** Work out these integrals by setting $u = \sec x$.

(a) $\int \tan^3 x \sec x\, dx.$ (b) $\int \sec^3 x \tan^5 x\, dx.$

**8.** If $\int \tan^m x \sec^n x\, dx$ is to work out easily by letting $u = \sec x$, no matter what kind of integer $n$ is, what appropriate condition should be put on $m$?

**9.** (a) Derive formula 26 of the Table of Integrals at the end of the book. *Suggestion*: write $\tan^n x = \tan^{n-2} x\,(\sec^2 x - 1) = \tan^{n-2} x \sec^2 x - \tan^{n-2} x$ and go on from there. Use a similar method to derive formula 27.
(b) Use formulas 26, 27 to work out $\int \tan^5 2x\, dx$ and $\int \operatorname{ctn}^6 3x\, dx$.

**10.** (a) Derive formula 28 of the Table of Integrals at the end of the book by setting $u = \sec^{n-2} x$, $dv = \sec^2 x\, dx$ and integrating by parts. In the $v\, du$ integral replace $\tan^2 x$ by $\sec^2 x - 1$ and go on from there. Use a similar method to derive formula 29.
(b) Use formulas 28, 29 to work out $\int \sec^5 4\theta\, d\theta$ and $\int \csc^3 ax\, dx$.

**11.** Use formulas 30 and 24, 25, 28, 29 of the Table of Integrals to work out:

(a) $\int \sin^4 x \cos^2 x\, dx.$ (c) $\int \dfrac{\cos^2 x}{\sin^4 x}\, dx.$

(b) $\int \sin^2 x \cos^6 x\, dx.$ (d) $\int \dfrac{\sin^4 x}{\cos^2 x}\, dx.$

**12.** Work out the following integrals by any method.

(a) $\int \sec^3 x\, dx.$ (d) $\int \frac{\sin^2 x}{\cos^6 x}\, dx.$

(b) $\int \tan^3 6x\, dx.$ (e) $\int \frac{\tan x}{\cos^2 x}\, dx.$

(c) $\int \operatorname{ctn}^2 x \sin x\, dx.$ (f) $\int \tan^2 x \sin^2 x\, dx.$

**13.** Proceed as directed in Exercise 12.

(a) $\int \frac{d\theta}{\sin^3 \theta}.$ (d) $\int \operatorname{ctn}^4 3x \csc^6 3x\, dx.$

(b) $\int \sec x \csc x\, dx.$ (e) $\int \frac{\operatorname{ctn} x}{\sin^4 x}\, dx.$

(c) $\int \tan^2 2x \sec^4 2x\, dx.$ (f) $\int \operatorname{ctn}^4 x \cos^2 x\, dx.$

**14.** Find the values of

(a) $\int_0^{\pi/2} \cos^5 x\, dx.$ (c) $\int_0^{\pi/2} \cos^{10} x\, dx.$

(b) $\int_0^{\pi/2} \sin^6 x\, dx.$ (d) $\int_0^{\pi/2} \sin^7 x\, dx.$

**15.** Work out the following integrals:

(a) $\int \sin x \cos 2x\, dx.$ (c) $\int \sin^2 x \cos 4x\, dx.$

(b) $\int \sin 3x \cos 4x\, dx.$ (d) $\int \cos^2 2x \sin 3x\, dx.$

**16.** Prove that, if $m$ and $n$ are positive integers, $\int_0^{2\pi} \sin mx \cos nx\, dx = 0,$ and that $\int_0^{2\pi} \sin mx \sin nx\, dx = \int_0^{2\pi} \cos mx \cos nx\, dx = 0$ if the further condition $m \neq n$ is satisfied.

**17.** (a) Make the substitution $y = \sin^{-1} x$ in the integral $\int x^n \sin^{-1} x\, dx$. In the resulting integral set $u = y$, $dv = \sin^n y \cos y\, dy$ and show that the integration can be finished with a reduction formula.
(b) Apply this method to the case $n = 3$.

**18.** (a) Develop a method, similar to that explained in Exercise 17, for $\int x^n \cos^{-1} x\, dx$.
(b) Apply the method to the case $n = 2$.

## 10-7 Trigonometric Substitutions

An integral of the form

$$\int x^m \sqrt{(a^2 - x^2)^n}\, dx, \tag{1}$$

where $m$ and $n$ are integers (they may be positive or negative) can be transformed in a useful way by the substitution $x = a \sin \theta$. Here we assume $a > 0$, and for convenience we assume that $\theta$ is an acute angle, so that all the trigonometric functions are positive. Then $dx = a \cos \theta \, d\theta$, and $a^2 - x^2 = a^2 - a^2 \sin^2 \theta = a^2 \cos^2 \theta$, so that (1) is transformed into

$$a^{m+n+1} \int \sin^m \theta \cos^{n+1} \theta \, d\theta. \tag{2}$$

This trigonometric integral is of the type considered in § 10-6. The best method for dealing with it will depend on the values of $m$ and $n$.

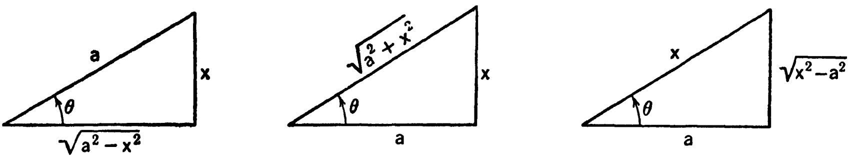



**Fig. 10-1**

Similar procedures apply if we have $a^2 + x^2$ or $x^2 - a^2$ in place of $a^2 - x^2$. A convenient scheme for showing the appropriate substitutions in the various cases is shown in Fig. 10-1.

**Example 1:** Work out $\int \frac{\sqrt{16 - x^2}}{x} dx$.

Here $x = 4 \sin \theta$, and the integral becomes

$$4 \int \frac{\cos^2 \theta}{\sin \theta} d\theta = 4 \int \frac{1 - \sin^2 \theta}{\sin \theta} d\theta = 4 \int (\csc \theta - \sin \theta) \, d\theta$$
$$= 4 \log |\csc \theta - \operatorname{ctn} \theta| + 4 \cos \theta + C.$$

To change back to $x$, refer to the first triangle in Fig. 10-1 (with $a = 4$). We see that

$$\csc \theta = \frac{4}{x}, \quad 4 \cos \theta = \sqrt{16 - x^2}, \quad \operatorname{ctn} \theta = \frac{\sqrt{16 - x^2}}{x}.$$

Hence (assuming $x > 0$)

$$\int \frac{\sqrt{16 - x^2}}{x} dx = 4 \log \frac{4 - \sqrt{16 - x^2}}{x} + \sqrt{16 - x^2} + C.$$

**Example 2:** Consider $\int \frac{dx}{x\sqrt{3 + x^2}}$.

Here we put $x = \sqrt{3} \tan \theta$, $dx = \sqrt{3} \sec^2 \theta \, d\theta$, and use the middle triangle in Fig. 10-1. We obtain

$$\frac{\sqrt{3}}{3} \int \frac{\sec^2 \theta}{\tan \theta \sec \theta} d\theta = \frac{\sqrt{3}}{3} \int \csc \theta \, d\theta = \frac{\sqrt{3}}{3} \log |\csc \theta - \operatorname{ctn} \theta| + C$$
$$= \frac{\sqrt{3}}{3} \log \frac{\sqrt{3 + x^2} - \sqrt{3}}{x} + C.$$

**Example 3:** Consider $\int \frac{dx}{x^3\sqrt{x^2 - a^2}}$.

Here we put $x = a \sec \theta$, $dx = a \sec \theta \tan \theta \, d\theta$, and use the third triangle in Fig. 10-1. We obtain

$$\frac{1}{a^3} \int \frac{\sec \theta \tan \theta}{\sec^3 \theta \tan \theta} d\theta = \frac{1}{a^3} \int \cos^2 \theta \, d\theta = \frac{1}{2a^3} (\theta + \sin \theta \cos \theta) + C.$$

Here we have used formula 23 from the Integral Table at the end of the book. Going back to $x$, we obtain

$$\int \frac{dx}{x^3\sqrt{x^2 - a^2}} = \frac{1}{2a^3} \cos^{-1} \frac{a}{x} + \frac{\sqrt{x^2 - a^2}}{2a^2x^2} + C.$$

This is valid if $x > 0$; if $x < 0$, we should have $\cos^{-1} (a/-x)$. Both cases can be combined by writing $\cos^{-1} (a/|x|)$.

## EXERCISES

**1.** Work out each of the following indefinite integrals. Check answers by using the Table of Integrals.

(a) $\int \frac{dx}{(a^2 + x^2)^{3/2}}$.

(b) $\int \frac{x^2}{\sqrt{25 - x^2}} dx$.

(c) $\int \frac{dx}{(x^2 - 16)^{3/2}}$.

(d) $\int \frac{\sqrt{25 - y^2}}{y^2} dy$.

(e) $\int \frac{dt}{t^2\sqrt{9 - t^2}}$.

(f) $\int x^3\sqrt{a^2 - x^2}\, dx$.

(g) $\int \frac{dx}{(16 + x^2)^2}$.

(h) $\int \frac{dx}{x^3\sqrt{x^2 + 4}}$. ✓

**2.** Proceed as directed in Exercise 1.

(a) $\int \frac{dx}{x\sqrt{x^2 - a^2}}$.

(b) $\int \frac{dx}{x^2\sqrt{x^2 + 9}}$.

(c) $\int \frac{dx}{(a^2 - x^2)^{3/2}}$.

(d) $\int \frac{x^2}{(a^2 - x^2)^{3/2}} dx$.

(e) $\int \frac{x^2}{\sqrt{x^2 + a^2}} dx$.

(f) $\int \frac{\sqrt{x^2 - 9}}{x^2} dx$.

(g) $\int \frac{u^2}{(u^2 + a^2)^{3/2}} du$.

(h) $\int \frac{\sqrt{a^2 + x^2}}{x} dx$.

**3.** Work out the first forms of formulas 18 and 19 in the Table of Integrals.

**4.** Work out both cases of formula 21 in the Table of Integrals. Use a reduction formula to integrate $\sec^3 \theta$.

**5.** Begin by completing the square, and then use the methods of this section.

(a) $\int \frac{dx}{(2ax - x^2)^{3/2}}$.

(b) $\int \frac{x^2}{\sqrt{x^2 + 2ax}} dx$.

6. Integrate by parts and then use a trigonometric substitution.

(a) $\int \frac{1}{x^3} \sin^{-1} \frac{x}{3}\, dx.$ (b) $\int x^2 \tan^{-1} \frac{x}{2}\, dx.$

7. Find the smaller area cut from the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ by the line $2bx + ay = 2ab$.

8. Let $L_1$ and $L_2$ be the lines $x = \pm c$, through the foci of the hyperbola $b^2x^2 - a^2y^2 = a^2b^2$. Let $L_3$ and $L_4$ be the lines $y = \pm b^2/a$. Let $A_1$ be the area between the two branches of the hyperbola and between $L_3$ and $L_4$. Let $A_1$ be the sum of the two areas bounded by the hyperbola and the lines $L_1$, $L_2$. Find the ratio $A_1/A_2$.

## 10-8 Rationalizing Substitutions

Indefinite integrals in which radicals are involved can sometimes be worked out by means of a substitution which transforms the problem into one of finding an indefinite integral of a rational function.

In order to describe the applicability of certain methods we must first pause to explain what is meant by speaking of *a rational function of two variables*, say $R(s, t)$. We call $R(s, t)$ a rational function of $s$ and $t$ if it is expressible as a quotient in which numerator and denominator are polynomials in $s$ and $t$. A polynomial in $s$ and $t$ is a sum of a finite number of terms of the form $cs^pt^q$, where $c$ is a numerical coefficient and $p$ and $q$ are nonnegative integers. The rational function is then defined whenever the polynomial in the denominator is not equal to zero.

Now consider a radical of the form $\sqrt[n]{a + bx}$, where $a$ and $b$ are constants and $n$ is an integer ($n \geq 2$), and consider indefinite integrals of the type

$$\int R(x, \sqrt[n]{a + bx})\, dx,$$

where $R$ is a rational function of $x$ and the radical. The substitution $u = (a + bx)^{1/n}$ will change this integral into one of the type considered in § 10-4.

**Example 1:** Work out $\int \frac{x}{\sqrt{x-1} + 2}\, dx.$

Here we set $\sqrt{x-1} = u$, $u^2 = x - 1$, $2u\, du = dx$. Our integral becomes

$$\int \frac{u^2+1}{u+2} 2u\, du = 2\int \left(u^2 - 2u + 5 - \frac{10}{u+2}\right) du$$

$$= \frac{2}{3}u^3 - 2u^2 + 10u - 20 \log (u + 2) + C$$

$$= \frac{2}{3}(x-1)^{3/2} - 2(x-1) + 10(x-1)^{1/2} - 20 \log (\sqrt{x-1} + 2) + C.$$

When we encounter radicals of the form $\sqrt[n]{a + bx^2}$ in an indefinite integral, the substitution $u = (a + bx^2)^{1/n}$ will lead us to a rational expression provided we can express the integral in the form

$$\int R(x^2, \sqrt[n]{a + bx^2}) \cdot x\, dx.$$

The important thing to notice here is that we have the combination $x\, dx$, and that the rest of the expression under the integral sign is a rational function of $x^2$ (note the exponent 2) and the radical. The combination $x\, dx$ is essential when we let $u = (a + bx^2)^{1/n}$.

**Example 2:** Work out $\int \frac{\sqrt{16 - x^2}}{x}\, dx$ by this method.

Here we have $16 - x^2 = u^2$, $-x\, dx = u\, du$, and our integral becomes

$$\int \frac{-u^2}{16 - u^2}\, du = \int \frac{(16 - u^2) - 16}{16 - u^2}\, du = \int \left(1 - \frac{16}{16 - u^2}\right) du$$

$$= u - \frac{16}{8} \log \frac{4 + u}{4 - u} + C = \sqrt{16 - x^2} - 2 \log \frac{4 + \sqrt{16 - x^2}}{4 - \sqrt{16 - x^2}} + C.$$

Here we have used formula 17 from the Table of Integrals. This problem was solved in a different way in Exercise 1, § 10-7. The student should show that the two answers are equivalent.

There is an entirely different class of problems in which a certain substitution brings each problem to the form of finding an indefinite integral of a rational function. This is the class of integrals of the form

$$\int R(\sin x, \cos x)\, dx.$$

This notation indicates that we have a rational function $R(s, t)$, with $s$ and $t$ replaced by $\sin x$ and $\cos x$, respectively. The systematic substitution here is

$$u = \tan \frac{x}{2}, \qquad x = 2 \tan^{-1} u, \qquad dx = \frac{2\, du}{1 + u^2}.$$

To express $\sin x$ and $\cos x$ in terms of $u$, we have

$$\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2} = 2 \tan \frac{x}{2} \cos^2 \frac{x}{2} = \frac{2 \tan \frac{x}{2}}{1 + \tan^2 \frac{x}{2}},$$

$$\cos x = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2} = \frac{1 - \tan^2 \frac{x}{2}}{1 + \tan^2 \frac{x}{2}},$$

so that

$$\sin x = \frac{2u}{1 + u^2}, \qquad \cos x = \frac{1 - u^2}{1 + u^2}.$$

**Example 3:** Work out $\int \frac{dx}{2 + 3\cos x}$.

Using the indicated substitution and the resulting formulas, the integral becomes

$$\int \frac{1}{2 + [3(1 - u^2)/(1 + u^2)]} \cdot \frac{2\,du}{1 + u^2} = \int \frac{2\,du}{5 - u^2}$$

$$= \frac{1}{\sqrt{5}} \log \left| \frac{\sqrt{5} + u}{\sqrt{5} - u} \right| + C = \frac{1}{\sqrt{5}} \log \left| \frac{\sqrt{5} + \tan \frac{x}{2}}{\sqrt{5} - \tan \frac{x}{2}} \right| + C.$$

This last method would have worked on many of the problems which were considered in § 10-6, but for most of those problems the earlier methods would prove more convenient than this new method.

## EXERCISES

**1.** Work out the indicated indefinite integrals by the methods of this section.

(a) $\int \frac{x\,dx}{\sqrt{3 + 4x}}$.

(b) $\int \frac{x^3\,dx}{(x^2 - a^2)^{3/2}}$.

(c) $\int \frac{2x - 1}{(3x - 5)^{2/3}}\,dx$.

(d) $\int \frac{2x - x^3}{\sqrt{a^2 + x^2}}\,dx$.

(e) $\int \frac{x^2}{\sqrt{x + 2}}\,dx$.

(f) $\int \frac{dx}{x - x^{1/3}}$.

(g) $\int \frac{dx}{x\sqrt{1 + 4x}}$.

(h) $\int \frac{dx}{x\sqrt{x^2 + a^2}}$.

(i) $\int \frac{dx}{x^3\sqrt{a^2 - x^2}}$.

(j) $\int \frac{dx}{3 + 2\cos x}$.

(k) $\int \frac{dx}{2 + \sin x - \cos x}$.

(l) $\int \frac{\sin x}{4 + \sin x}\,dx$.

**2.** Proceed as directed in Exercise 1.

(a) $\int x\sqrt{2 - 5x}\,dx$.

(b) $\int (3x - 4x^3)\sqrt{9 - x^2}\,dx$.

(c) $\int \frac{\sqrt{1 - x}}{x}\,dx$.

(d) $\int x^3\sqrt{16 + 5x^2}\,dx$.

(e) $\int \frac{dx}{x^2\sqrt{4 - 3x}}$.

(f) $\int \frac{\sqrt{x + 2} + 3}{\sqrt{x + 2} - 3}\,dx$.

(g) $\int \frac{dx}{x\sqrt{a^2 - x^2}}$.

(h) $\int \frac{dx}{x^3\sqrt{a^2 + x^2}}$.

(i) $\int \frac{dx}{5 - 4\cos x}$.

(j) $\int \frac{dx}{\sin x - \cos x - 1}$.

(k) $\int \frac{dx}{2 - 3\sin x}$.

(l) $\int \frac{\cos x}{5 - 2\cos x}\,dx$.

## 10-9 Substitution and Change of Limits

If a definite integral is evaluated by making a substitution of some kind, it is possible to express the integral as a definite integral with respect to the new variable, the limits being those values of the new variable which correspond to the original limits of integration. We shall give an illustration, and then make the situation precise in a theorem.

**Example 1:** Find the value of $\int_{-3}^{6} x^3\sqrt{36 - x^2}\,dx$.

We let $u = \sqrt{36 - x^2}$, so that $u^2 = 36 - x^2$ and $u\,du = -x\,dx$. When $x = -3$, $u = 3\sqrt{3}$, and when $x = 6$, $u = 0$. Hence

$$\int_{-3}^{6} x^3\sqrt{36 - x^2}\,dx = \int_{3\sqrt{3}}^{0} (36 - u^2)u(-u\,du)$$

$$= \int_{0}^{3\sqrt{3}} (36u^2 - u^4)\,du = \frac{2673\sqrt{3}}{5}.$$

Note that we considered $x^3\,dx$ as $x^2 \cdot x\,dx$. The last calculations are left to the student.

Here is a general theorem which covers this procedure.

THEOREM 10-A. *Let $f(x)$ be continuous when $a \leq x \leq b$. Suppose that $x$ is set equal to a function of a new variable $u$ which ranges over an interval $[\alpha, \beta]$, and suppose the following conditions are satisfied:*

(i) *$dx/du$ is continuous when $\alpha \leq u \leq \beta$;*
(ii) *$x$ lies in $[a, b]$ when $u$ lies in $[\alpha, \beta]$;*
(iii) *$x = a$ when $u = \alpha$ and $x = b$ when $u = \beta$.*

*Suppose, finally, that the change of variable transforms $f(x)\,dx$ into $\phi(u)\,du$. Then*

$$\int_{a}^{b} f(x)\,dx = \int_{\alpha}^{\beta} \phi(u)\,du. \tag{1}$$

*Proof*. Form the integrals

$$F(x) = \int_{a}^{x} f(s)\,ds, \qquad \Phi(u) = \int_{\alpha}^{u} \phi(t)\,dt.$$

In this notation, to prove (1) is the same as proving that $F(b) = \Phi(\beta)$. Now we know by Theorem 6-C that

$$F'(x) = f(x) \quad \text{and} \quad \Phi'(u) = \phi(u). \tag{2}$$

Let the dependence of $x$ on $u$ be expressed as $x = g(u)$. Then $dx = g'(u)\,du$, and

$$f(x)\,dx = f[g(u)]g'(u)\,du,$$

and so $f(x)\,dx = \phi(u)\,du$ means that

$$\phi(u) = f[g(u)]g'(u). \tag{3}$$

Now consider $F[g(u)]$. Its derivative with respect to $u$ is $F'[g(u)]g'(u)$. By (3) and (2) we see that

$$\frac{d}{du} F[g(u)] = \Phi'(u).$$

Hence $F[g(u)] = \Phi(u) + C$, where $C$ is some constant. If we put $u = \alpha$, then $x = g(\alpha) = a$, and $F(a) = \Phi(\alpha) + C$. Since $F(a)$ and $\Phi(\alpha)$ are both zero, we see that $C = 0$. Now put $u = \beta$, with the result $F[g(\beta)] = \Phi(\beta)$. Since $g(\beta) = b$, we have finished the proof in the manner stated at the outset.

In stating the theorem we indicated that $a < b$ and $\alpha < \beta$. But it would make no difference if $a < b$ and $\beta < \alpha$, as long as $x = a$ corresponds to $u = \alpha$, and likewise for the other limits. This is the way it was in the illustrative example.

In some problems the change of limits is especially advantageous because of the possibility it offers of using conveniently tabulated integrals.

**Example 2:** Calculate $\int_0^a (a^2 - x^2)^{n/2}\, dx$, where $a > 0$ and $n$ is an odd positive integer, say $n = 2p - 1$, where $p \geq 1$.

Letting $x = a \sin \theta$, where $0 \leq \theta \leq \pi/2$), we obtain

$$\int_0^{\pi/2} (a \cos \theta)^n\, a \cos \theta\, d\theta = a^{2p} \int_0^{\pi/2} \cos^{2p} \theta\, d\theta.$$

Now we can use formula 107 from the Table of Integrals. The value of our integral is

$$a^{2p}\frac{1\cdot 3 \cdots (2p-1)}{2\cdot 4 \cdots 2p}\cdot\frac{\pi}{2} = \frac{(2p)!}{(2^p p!)^2}\cdot\frac{\pi}{2}a^{2p}.$$

## EXERCISES

**1.** Calculate each of these definite integrals by making a substitution and a corresponding change in the limits of integration.

(a) $\int_5^{13} \frac{\sqrt{x-4}}{x^2-16}\, dx.$

(b) $\int_0^4 \frac{dx}{(16+x^2)^2}.$

(c) $\int_1^{14} \frac{x^2\, dx}{(2x-1)^{2/3}}.$

(d) $\int_0^4 (16 - x^2)^3\, dx.$

(e) $\int_0^a x^4\sqrt{a^2 - x^2}\, dx.$

(f) $\int_0^8 (4 - x^{2/3})^{5/2}\, dx.$

**2.** Use half-angle and double-angle formulas and the Integral Table formulas No. 107 and No. 108 to calculate the following integrals.

(a) $\int_0^{\pi} (1 - \cos \theta)^{5/2}\, d\theta.$

(b) $\int_0^{\pi} (1 + \cos \theta)^{7/2}\, d\theta.$

(c) $\int_0^{\pi/4} \sin^6 \theta\, d\theta.$

(d) $\int_0^{\pi/4} \cos^8 \theta\, d\theta.$

## 10-10 Tables of Integrals

In this chapter we have illustrated how a certain amount of system can be brought into the business of finding antiderivatives of specified functions. To make it easier to calculate integrals as they arise in practice, many indefinite integrals (that is, antiderivatives) have been tabulated so that they may be referred to as needed. A number of tables of this kind are available in various mathematical handbooks. A small table of indefinite integrals, adequate for most of the problems the student will find in this book, is contained in the book itself, at the back.

In order to be able to use a table of integrals to the best advantage, the student must study the arrangement of the tables, observing the manner in which the integrals are classified. He must also be able to perform any preliminary transformations or simplifications which may be necessary to bring a given integral into a form which is tabulated.

**Example 1:** Evaluate the integral $\int xe^{-\sqrt{x}}\,dx$.

We make the substitution $y^2 = x$. The integral then becomes

$$2\int y^3 e^{-y}\,dy.$$

This is dealt with by means of the reduction formula (see Table of Integrals, No. 85)

$$\int y^n e^{ay}\,dy = \frac{1}{a}y^n e^{ay} - \frac{n}{a}\int y^{n-1}e^{ay}\,dy. \tag{1}$$

The final result is

$$\int y^3 e^{-y}\,dy = -e^{-y}(y^3 + 3y^2 + 6y + 6) + C.$$

The original integral, therefore, has the value

$$\int xe^{-\sqrt{x}}\,dx = -2e^{-\sqrt{x}}(x^{3/2} + 3x + 6x^{1/2} + 6) + C.$$

**Example 2:** Evaluate the integral

$$I = \int_0^{\pi/2} \frac{\cos^2 x\,dx}{1 + \cos^2 x}.$$

We first make the algebraic reduction

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

Next, introduce the trigonometric identity which expresses $\cos^2 x$ in terms of $\cos 2x$:

$$\frac{1}{1 + \cos^2 x} = \frac{1}{1 + \dfrac{1 + \cos 2x}{2}} = \frac{2}{3 + \cos 2x}.$$

Thus, $$I = \int_0^{\pi/2} \left(1 - \frac{2}{3 + \cos 2x}\right) dx = \frac{\pi}{2} - \int_0^{\pi/2} \frac{2\,dx}{3 + \cos 2x}.$$

The general formula which is needed here is (Table of Integrals, No. 95)

$$\int \frac{du}{a + b\cos u} = \frac{2}{\sqrt{a^2 - b^2}} \tan^{-1} \frac{\sqrt{a^2 - b^2}\tan \frac{1}{2}u}{a + b} + C. \tag{2}$$

This is valid if $a^2 - b^2 > 0$. To use it, we must set $a = 3$, $b = 1$, $u = 2x$. Then,

$$\int_0^{\pi/2} \frac{2\,dx}{3 + \cos 2x} = \frac{2}{2\sqrt{2}} \tan^{-1} \frac{2\sqrt{2}\tan x}{4}\bigg|_0^{\pi/2} = \frac{1}{\sqrt{2}}\left(\frac{\pi}{2} - 0\right).$$

Thus, finally, $$I = \frac{\pi}{2}\left(1 - \frac{1}{\sqrt{2}}\right).$$

*CHAPTER* XI

# FURTHER APPLICATIONS OF INTEGRATION

## 11-1 Arc Length

In elementary geometry, the circumference of a circle is found as the limit of the perimeters of regular polygons inscribed in the circle. This is a particular instance of the general procedure for defining the length of an arc of a curve.

Let $C$ be a given curved arc, with end points $A$, $B$. Let us insert points in order along $C$ from $A$ to $B$, so that $C$ is divided into $n$ pieces. Let these points be $P_0, P_1, \cdots, P_n$, with $P_0 = A$, $P_n = B$ (see Fig. 11-1). Then let us draw the line segments joining $P_0$ to $P_1$, $P_1$ to $P_2$, and so on. It seems intuitively plausible to consider the sum of the lengths of these segments as an approximate measure of the length of $C$. This is, in fact, the way we propose to *define* the length of $C$. There is no clear mathematical meaning for the length of a curve until we have made the meaning clear by a definition. And just as in the case of defining the area of a plane figure with a curved boundary, the length of a curve must be defined by some kind of a limiting process, starting from the simple things whose length we *do* know, namely, line segments.

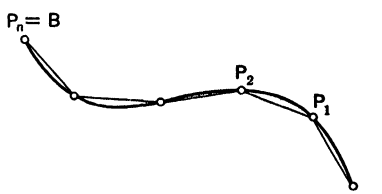



Fig. 11-1

We shall define the length $L$ of $C$ as the limit of the sum

$$\overline{P_0P_1} + \overline{P_1P_2} + \cdots + \overline{P_{n-1}P_n}, \tag{1}$$

provided that we can show that this sum does indeed approach a limit as the number $n$ is increased and the greatest length of the individual segments $P_0P_1$, $P_1P_2$, $\cdots$ is made to approach zero. In order to show that the sums do approach a limit we must have some rather exact information about the nature of the curve $C$.

We therefore begin by considering a case of general interest in which we can accomplish this goal. We suppose that $C$ is the graph of $y = f(x)$, where $f$ is a function which has a continuous derivative, and $x$ varies from $a$ to $b$, where $a < b$. In this case we shall show that the length of $C$ is

$$L = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx. \tag{2}$$

Hence, to calculate $L$, we merely work out the value of the integral.

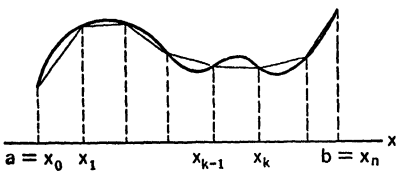


**Fig. 11-2**

In order to derive the formula (2), consider Fig. 11-2. Here the points $P_0, \cdots, P_n$ along $C$ have been determined by choosing points $x_0, x_1, \cdots, x_n$ along the $x$-axis from $a$ to $b$. If $y_k = f(x_k)$, then $P_k$ is the point $(x_k, y_k)$. Now

$$\overline{P_{k-1}P_k} = [(x_k - x_{k-1})^2 + (y_k - y_{k-1})^2]^{1/2}. \tag{3}$$

Since $f$ is continuous, it is evident from (3) that $\overline{P_{k-1}P_k} \to 0$ if $(x_k - x_{k-1}) \to 0$. And certainly the reverse implication is true, because $x_k - x_{k-1} \leq \overline{P_{k-1}P_k}$. Hence, in this case, we have to find the limit of the sum (1) as the greatest of the differences $x_k - x_{k-1}$ approaches zero.

Now let us simplify (3) by using the law of the mean (Theorem 2-C). There is some number $X_k$ between $x_{k-1}$ and $x_k$ such that

$$y_k - y_{k-1} = f(x_k) - f(x_{k-1}) = (x_k - x_{k-1})f'(X_k).$$

We write $\Delta x_k = x_k - x_{k-1}$ for convenience. Then (3) becomes

$$\overline{P_{k-1}P_k} = \{1 + [f'(X_k)]^2\}^{1/2}\, \Delta x_k.$$

But the limit of the sum of all these things is exactly the integral (2), by the definition of the integral.

**Example 1:** Find the length of the arc of the parabola $4y = x^2$ from $(-2, 1)$ to $(4, 4)$.

Here $dy/dx = x/2$, so the formula is

$$L = \int_{-2}^{4} \sqrt{1 + \frac{x^2}{4}}\, dx = \frac{1}{2}\int_{-2}^{4} \sqrt{4 + x^2}\, dx.$$

$$= \left[\frac{x}{4}\sqrt{4 + x^2} + \log\left(x + \sqrt{4 + x^2}\right)\right]_{-2}^{4}$$

$$= \sqrt{20} + \log(4 + \sqrt{20}) + \frac{1}{2}\sqrt{8} - \log(-2 + \sqrt{8}).$$

This can be reduced to

$$L = 2\sqrt{5} + \sqrt{2} + \log \frac{\sqrt{5} + 2}{\sqrt{2} - 1}.$$

If we use the inverse hyperbolic sine form of the indefinite integral [see (8) in § 9-3], the answer is

$$L = 2\sqrt{5} + \sqrt{2} + \sinh^{-1} 2 + \sinh^{-1} 1.$$

For some purposes it is convenient to deal with the length of arc from $A$ to a variable point $P$ moving along the curve $C$. If $A$ corresponds to $x = a$ and $P$ corresponds to a variable value of $x$, then the arc length $s$ from $A$ to $P$ is

$$s = \int_a^x \sqrt{1 + [f'(t)]^2}\, dt. \tag{4}$$

Here we have used $t$ instead of $x$ as a variable of integration, because $x$ is being used for another purpose. If we regard $s$ as a function of $x$ defined by (4), the fundamental theorem about derivatives and integrals (Theorem 6-C) tells us that

$$\frac{ds}{dx} = \sqrt{1 + [f'(x)]^2}. \tag{5}$$

This is often written in the alternative forms

$$\left(\frac{ds}{dx}\right)^2 = 1 + \left(\frac{dy}{dx}\right)^2, \quad \text{or} \quad ds^2 = dx^2 + dy^2. \tag{6}$$

This formula relating $ds$ to $dx$ and $dy$ will be used in studying the *curvature* of curves and in studying the motion of particles in curved paths (see Chapter XIII).

It is also important to know how to find the length of an arc of a curve if it is represented parametrically, as in § 5-7. Let us suppose that the parametric equations are

$$x = \phi(t), \qquad y = \psi(t),$$

where $t$ varies from $a$ to $b$ and $\phi$, $\psi$ have continuous derivatives which are never zero for the same value of $t$. These conditions have the effect of making the curve smooth, with a tangent whose inclination is a continuous function of $t$. Also, if $P$ and $P'$ are points on the curve corresponding to $t$ and $t'$, respectively, then $\overline{PP'} \to 0$ is equivalent to $|t - t'| \to 0$, provided we assume the curve does not intersect itself. The length of the arc which is generated as $t$ varies from $a$ to $b$ is now

$$L = \int_a^b \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2\right]^{1/2} dt. \tag{7}$$

The proof of this formula is somewhat more complicated than was the derivation of (2). A closer analysis of the situation is made in § 11-3.

Meanwhile we go ahead and use the formula. In this case the differential formula is

$$\frac{ds}{dt} = \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2\right]^{1/2}, \quad \text{or} \quad ds^2 = dx^2 + dy^2. \tag{8}$$

**Example 2:** Find the length of one arch of the cycloid (see § 5-8). The parametric equations are

$$x = a(\theta - \sin\theta), \qquad y = a(1 - \cos\theta).$$

For one arch, the parameter $\theta$ goes from 0 to $2\pi$. Now

$$\frac{dx}{d\theta} = a(1 - \cos\theta), \qquad \frac{dy}{d\theta} = a\sin\theta,$$

and

$$\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2 = a^2(1 - \cos\theta)^2 + a^2\sin^2\theta$$

$$= 2a^2(1 - \cos\theta) = 4a^2\sin^2\frac{\theta}{2}.$$

Therefore the required length is

$$L = 2a\int_0^{2\pi}\sin\frac{\theta}{2}\,d\theta = -4a\cos\frac{\theta}{2}\Big]_0^{2\pi} = 8a.$$

In practice the student may prefer to condense his basic information about arc lengths into the form

$$\text{arc length} = \int ds, \qquad ds^2 = dx^2 + dy^2.$$

Then $dx$ and $dy$ may be calculated in terms of whatever independent variable is the most convenient (and, of course, the differential of that variable).

It often turns out that integrals expressing arc length are not easy to evaluate, for the reason that the function under the integral sign does not have any elementary function as an antiderivative. This occurs with the ellipse; the nonelementary integral in this case is called an elliptic integral (see Exercise 5).

## EXERCISES

**1.** Find the arc length of each curve between the points indicated.
   (a) $y = x^{3/2}$ from $x = 0$ to $x = 4$;
   (b) $y = \log x$ from $x = \frac{1}{2}$ to $x = 2$;
   (c) $y = \log\cos x$ from $x = 0$ to $x = \pi/3$;
   (d) $y = \log(1 - x^2)$ from $x = 0$ to $x = \frac{3}{4}$;
   (e) $y = \frac{1}{2}(e^x + e^{-x})$ from $x = -1$ to $x = 1$;
   (f) $(y + 1)^2 = 4x^3$ from $(0, -1)$ to $(1, 1)$;
   (g) $x^2 + 2y + 2 = 0$ from $(-\sqrt{2}, -2)$ to $(0, -1)$;
   (h) $y = \dfrac{x^2}{2} - \dfrac{1}{4}\log x$ from $x = 1$ to $x = 2$.

**2.** Find the arc length of each curve between the points indicated. Integrate with respect to $y$.
   (a) $y^2 = -4x$ from $(-4, 4)$ to $(0, 0)$;

(b) $3x^2 = y^3$ from $(-3, 3)$ to $(8/\sqrt{3}, 4)$;
(c) the shorter arc of $x^2 + y^2 = 32$ from $(4, 4)$ to $(2\sqrt{6}, -2\sqrt{2})$;
(d) $y = e^{-2x}$ from $y = \frac{1}{4}$ to $y = 4$;
(e) $y = \sin^{-1}(e^x)$ from $y = \pi/6$ to $y = \pi/2$.

**3.** Find the arc length of each curve between the points indicated.
(a) $x = 2t^2$, $y = t^3$ from $t = 1$ to $t = 2$;
(b) $x = \frac{1}{2}\log(t^2 - 1)$, $y = \sqrt{t^2 - 1}$ from $t = 3$ to $t = 7$;
(c) $x = 5\sin t$, $y = 5\cos t$ from $t = -\pi/3$ to $t = \pi/2$;
(d) $x = 4 + 2t$, $y = \frac{1}{2}t^2 + 3$ from $t = -2$ to $t = 2$;
(e) $x = e^t\cos t$, $y = e^t\sin t$ from $t = 0$ to $t = 2$;
(f) $x = 2\cos^2\theta + \sin 2\theta$, $y = \sin 2\theta + 2\sin^2\theta$ from $\theta = -\pi/4$ to $\theta = 3\pi/4$;
(g) $2x = e^t - e^{-t}$, $8y = e^{2t} + e^{-2t} - 4$, from $t = 0$ to $t = \log(3 + 2\sqrt{2})$;
(h) $x = \sqrt{a^2 - t^2}$, $y = a\log\dfrac{a + t}{\sqrt{a^2 - t^2}} - t$ from $t = 0$ to $t = \dfrac{a\sqrt{3}}{2}$.

**4.** Find the length of the curve $x = 9t^2$, $y = 9t^3 - 3t$ corresponding to $0 \le t \le 1/\sqrt{3}$.

**5.** Show that the total perimeter of the ellipse $9x^2 + 25y^2 = 225$ can be expressed in either of the forms

$$4\int_0^{\pi/2}\sqrt{9 + 16\sin^2\theta}\,d\theta \quad\text{or}\quad 4\int_0^{\pi/2}\sqrt{25 - 16\cos^2\theta}\,d\theta.$$

Use the parametrization $x = 5\cos\theta$, $y = 3\sin\theta$. What would be the result if we wrote $x = 5\sin t$, $y = 3\cos t$?

**6.** The following arc-length problems lead to nonelementary integrals. Set up definite integrals for each case.
(a) The arc of $xy = 1$ from $x = 1$ to $x = 4$.
(b) The arc of $y = x^3$ from $x = 0$ to $x = 1$.
(c) The arc of $x^2 - y^2 = 1$ from $(\sqrt{2}, -1)$ to $(\sqrt{10}, -3)$.

**7.** Consider the circle $x^2 + y^2 = 1$. Let $P_1$ and $P_2$ be on the first quadrant arc of the circle, with $y$-coordinates $y_1$, $y_2$ such that $0 \le y_1 < y_2 < 1$. Draw a figure. Express the arc length from $P_1$ to $P_2$ as an integral with respect to $y$ and evaluate it, thus verifying that this method gives results consistent with the use of radian measure for angles. It would be logically permissible to *define* radian measure by using this arc-length integral, and then to go on to develop trigonometry from this starting point.

## 11-2 Solids of Revolution: Shell Method

We shall now describe a second method for finding the volume of a solid of revolution. For the first method see the discussion leading up to (9) in § 6-1. Also, see § 6-7. Suppose that the volume is generated by revolving about the $y$-axis an area lying all on one side of the $y$-axis in the $xy$-plane. Let the area in question extend from $x = a$ to $x = b$, and suppose that the area is bounded above and below by curves whose equations are known. If the area is divided into narrow strips parallel to the $y$-axis, a typical strip of height $h(x)$ (a known function of $x$) and width $\Delta x$ will generate a thin-walled cylindrical shell (see Fig. 11-3). The area of the inner surface of

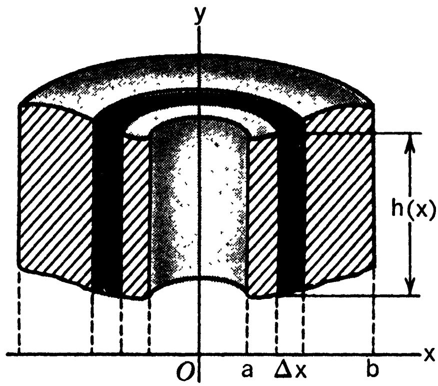



**Fig. 11-3**

this shell is $2\pi x h(x)$; the area of the outer surface is a similar expression with $x$ replaced by $x + \Delta x$. In order to approximate the volume of this shell we imagine it to be split open and unrolled so as to form a thin rectangular sheet of area $2\pi x h(x)$ and thickness $\Delta x$. In this way we are led to the expression

$$2\pi x h(x)\, \Delta x \tag{1}$$

as an approximation to the volume of the shell. The limit of the sum of the expressions (1) as the maximum $\Delta x$ approaches zero is the integral

$$2\pi \int_a^b x h(x)\, dx. \tag{2}$$

Hence it is plausible to accept this integral as the correct expression for the entire volume of the solid of revolution.

**Example:** The area above the curve $8y = 12x - x^3$ and below the line $y = 2$, from $x = 0$ to $x = 2$, is revolved about the $y$-axis. Find the volume generated (see Fig. 11-4).

Here the height of a typical strip is $2 - y$, where $y$ is found from the equation of the curve. When the strip is revolved about the $y$-axis it generates a cylindrical shell of altitude $2 - y$, inner radius $x$, and outer radius $x + \Delta x$. The approximation to the volume of the shell is $2\pi x(2 - y)\, \Delta x$, and the total volume under consideration is

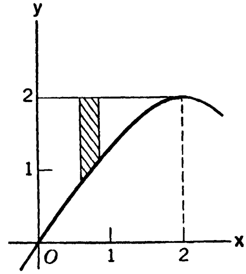



**Fig. 11-4**

$$V = 2\pi \int_0^2 x(2 - y)\, dx.$$

Now,
$$2 - y = 2 - \frac{3}{2}x + \frac{1}{8}x^3,$$

and so $$V = 2\pi \int_0^2 \left(2x - \frac{3}{2}x^2 + \frac{1}{8}x^4\right) dx = \frac{8\pi}{5}.$$

The details of the integration are left to the student.

We now have two different methods for finding the volume of a solid by integration. The first method is that of slicing the solid into thin parallel plane sections. The volume of each slice is approximately the product of the thickness of the slice and the area of the section. The second method is that of thin cylindrical shells. The volume of a shell is approximately the product of the thickness of the shell and the lateral area of the shell.

Now, we *defined* the volume of a solid as the value of the integral arrived at by the method of slicing into plane sections. The method of cylindrical shells has furnished us a different integral formula for finding volumes. Our derivation of this formula was not based on the definition of volume in terms of plane sections, but was merely supported by an argument of what seems to be plausible in view of our intuitive notions about volume. Logically, then, we still lack a rigorous proof that the shell method is consistent with the slicing method; that is, we have not proved that when the two methods are applied to the same problem, they will give the same answer. Such a proof is best deferred until we study double integrals; see § 20-3, Exercise 14. For another critique of the shell method see Example 1, § 11-3.

## EXERCISES

**1.** Find the volume of the right circular cone in Fig. 11-5 by the shell method.

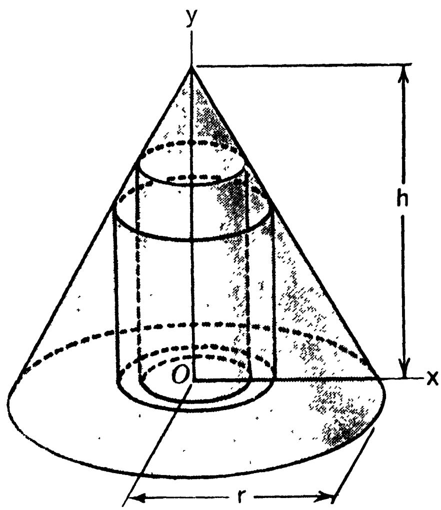



**Fig. 11-5**

**2.** Find the volume of the paraboloid in Fig. 11-6 by the shell method.

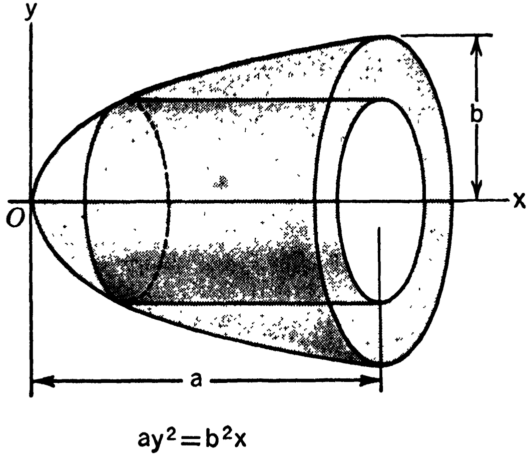



**Fig. 11-6**

**3.** The ellipse $b^2x^2 + a^2y^2 = a^2b^2$ is revolved about the $y$-axis, generating a spheroid. Find the volume of the spheroid by the shell method.

**4.** The line $y = x \operatorname{ctn} \alpha$ (where $0 < x < \pi/2$) and the circle $x^2 + y^2 = a^2$ are simultaneously revolved about the $y$-axis. Thus a sphere and a full cone (two nappes) are generated. Find the volume which is both inside the sphere and inside one nappe of the cone.

**5.** In each part of this exercise an area is described. Find the volume generated when it is revolved about the line indicated.

(a) The area in the first quadrant, between $xy = 64$ and $x + y = 20$, about the $x$-axis.

(b) The area between the $x$-axis and $y(4 + x^2) = 16$, from $x = 0$ to $x = 2$, about the $y$-axis.

(c) The area between $x^2 = 4y$ and $x^2 + 4 = 8y$, about the $y$-axis.

(d) The area between the $y$-axis and $y^2 + \log x = 0$, from $y = 0$ to $y = \sqrt{2}/2$, about the $x$-axis.

(e) The area between $y = e^x$ and the $x$-axis, from $x = 0$ to $x = 1$, about the $y$-axis.

(f) The area under the arch of the curve $y = 4 \sin 2x$, $0 \le x \le \pi/2$, about the $y$-axis.

(g) The area bounded by the hyperbola $16y^2 - 9x^2 = 144$ and the line $y = 6$, about the $x$-axis.

(h) The area in (g), about the $y$-axis.

(i) The area in the first quadrant bounded by $y = x^2$, $x = 0$, $y = 1$, about the line $x = 1$.

(j) The area in the first quadrant bounded by $4y^2 = x$, $x = 0$, $y = 1$, about the line $y = 2$.

(k) The area bounded by the parabola $2x^2 = y$ and the line $2x - y + 4 = 0$, about the line $x = 2$.

**6.** Find the volume of the torus generated by revolving a circle of radius $a$ about a line in its plane whose distance from the center is $b$, where $b > a$.

## 11-3 The Principle of Duhamel

In the applications of integral calculus the typical procedure consists in formulating geometrical or physical quantities as limits of sums of appropriately selected small parts or *elements*. In each case the limit of the sum is recognized as a definite integral. Sometimes we recognize the limit of the sum as an integral by the very definition of an integral. In other cases the recognition calls for mathematical justification.

Let us introduce some terminology and notation to help us in our discussion. When an interval $a \leqq x \leqq b$ is divided into subintervals, let us say that we form a *partition* of the interval, and let us use a symbol $\Delta$ for such a partition. A partition is formed by choosing any finite number of points $x_1, x_2, \cdots, x_{n-1}$ such that $a < x_1 < x_2 < \cdots < x_{n-1} < b$. We then write $x_0 = a$, $x_n = b$. There are $n$ subintervals, of lengths $\Delta x_1 = x_1 - x_0$, $\Delta x_2 = x_2 - x_1, \cdots, \Delta x_n = x_n - x_{n-1}$. The largest of these lengths is called the *norm* of the partition, and denoted by $\|\Delta\|$. Observe that the partition is not determined merely by the number of its subintervals, but by the distribution of the points $x_1, \cdots, x_{n-1}$. Note also that if $\|\Delta\|$ is made very small in comparison with $b - a$, then $n$ must of necessity become very large.

We form partitions of an interval each time that we undertake to use the methods of integral calculus to derive a new definite integral formula for some geometrical or physical quantity. The variable under consideration is usually a coordinate of some kind, or a parameter related to the problem under discussion. In our work just now let us denote this variable by $x$, and let us suppose the interval is $a \leqq x \leqq b$. By forming a partition of the interval, we divide the physical or geometrical quantity into parts, one part corresponding to each subinterval in the partition. Let $Q$ denote the numerical quantity we wish to compute. For example, $Q$ might be an area, a volume, an arc length, a moment of inertia, or a component of gravitational attraction in some problem. If $\Delta$ is the partition, let $\Delta Q_1$, $\Delta Q_2, \cdots, \Delta Q_n$ denote either the exact parts of $Q$ corresponding to the subintervals of lengths $\Delta x_1, \cdots, \Delta x_n$, so that

$$Q = \Delta Q_1 + \Delta Q_2 + \cdots + \Delta Q_n, \tag{1}$$

or else let them denote approximations to the parts of $Q$, of such a nature that $Q$ is the limit of the sum:

$$Q = \lim (\Delta Q_1 + \Delta Q_2 + \cdots + \Delta Q_n) \tag{2}$$

as $\|\Delta\| \to 0$.

As an illustration, let $Q$ be the volume of the solid of revolution considered in § 11-2, and let $\Delta Q$ denote the exact volume of the shell of wall thickness $\Delta x$ shown in Fig. 11-3. Then $Q$ is the exact sum of all such $\Delta Q$ when we make a partition of the interval $[a, b]$. As another illustration, let $Q$ be the length of arc in Fig. 11-2, and let $\Delta Q_i$ be the chord length $\overline{P_{i-1}P_i}$. Here $\Delta Q_i$ is not the exact arc length $P_{i-1}P_i$, but (2) is true by definition in this case.

The next step in the general process is to attempt to find a continuous function $F(x)$ such that each $\Delta Q_i$ is either exactly or approximately equal to $F(x_i') \Delta x_i$, where $x_i'$ is some point between $x_{i-1}$ and $x_i$. If $\Delta Q_i = F(x_i') \Delta x_i$ exactly, then

$$\Delta Q_1 + \cdots + \Delta Q_n = F(x_1') \Delta x_1 + \cdots + F(x_n') \Delta x_n. \tag{3}$$

The sum on the right in (3) approaches the definite integral of $F(x)$ as its limit when $\|\Delta\| \to 0$. Therefore

$$Q = \int_a^b F(x)\, dx. \tag{4}$$

In practice it is not always easy to find a function $F(x)$ for which it is quickly apparent that $\Delta Q_i = F(x_i') \Delta x_i$ exactly. Suppose, however, that it is possible to find *two* continuous functions $f(x)$ and $\phi(x)$, and *two* points $x_i'$ and $x_i''$ between $x_{i-1}$ and $x_i$ such that $\Delta Q_i = f(x_i')\phi(x_i'') \Delta x_i$. Then, if $\Delta x_i$ is small, and if we define $F(x) = f(x)\phi(x)$, $\Delta Q_i$ is *approximately* equal to $F(x_i') \Delta x_i$. Now the limit formula

$$\lim_{\|\Delta\|\to 0} [f(x_1')\phi(x_1'') \Delta x_1 + \cdots + f(x_n')\phi(x_n'') \Delta x_n] = \int_a^b f(x)\phi(x)\, dx \tag{5}$$

is true, though its truth is not just a matter of definition. As a consequence of this formula we see that

$$Q = \int_a^b f(x)\phi(x)\, dx$$

provided $\Delta Q_i = f(x_i')\phi(x_i'') \Delta x_i$.

We omit the proof of (5), but we shall frequently use the formula itself. *The formula is a theorem about integrals.* The late Professor G. A. Bliss emphasized the usefulness of this theorem in calculus. We shall therefore refer to (5) as *the formula of Bliss.*

**Example 1:** Consider the derivation of the formula for the volume of a solid of revolution by the shell method, in §-11-2. For simplicity assume that the volume is generated by revolving about the $y$-axis the area under a curve $y = f(x)$, from $x = a$ to $x = b$, where $f(x)$ is positive and continuous. If we make a partition of the interval $[a, b]$, let $\Delta V_i$ be the volume of the shell generated by revolving the strip of area between $x_{i-1}$ and $x_i$. Let $m_i$ and $M_i$ be the minimum and maximum values of $f(x)$ between these values of $x$. Then the volume $\Delta V_i$ is certainly at least as great as the volume generated by revolving the strip if it were cut off at a height $m_i$, and no greater than the volume generated by the strip if it were of uniform height $M_i$. In this way we see, by considering concentric cylinders of radii $x_{i-1}$ and $x_i$, that

$$\pi(x_i^2 - x_{i-1}^2)m_i \leqq \Delta V_i \leqq \pi(x_i^2 - x_{i-1}^2)M_i.$$

Therefore, for some intermediate value of $f(x)$ at a point $x_i'$ between $x_{i-1}$ and $x_i$,

$$\Delta V_i = \pi(x_i^2 - x_{i-1}^2)f(x_i').$$

Now let

$$x_i'' = \tfrac{1}{2}(x_i + x_{i-1}),$$

and note that $\Delta x_i = x_i - x_{i-1}$. Then

$$\Delta V_i = 2\pi x_i'' f(x_i')\,\Delta x_i.$$

If we now set $\phi(x) = 2\pi x$, we see that

$$\Delta V_i = f(x_i')\phi(x_i'')\,\Delta x_i.$$

It then follows by the formula of Bliss that

$$V = 2\pi \int_a^b xf(x)\,dx.$$

This justifies the shell method of calculating volumes.

Let us now return to the general problem of trying to find a function $F(x)$ such that

$$\lim_{\|\Delta\|\to 0} (\Delta Q_1 + \cdots + \Delta Q_n) = \int_a^b F(x)\,dx. \tag{6}$$

The "practical" attitude which is adopted by many people in working with calculus is something like this: if one can find a continuous function $F(x)$ such that each $\Delta Q_i$ is approximately equal to $F(x_i')\,\Delta x_i$, then (6) holds. This is a good working principle, but its validity depends on a more exact definition of what is meant by saying that $\Delta Q_i$ and $F(x_i')\,\Delta x_i$ are approximately equal. The effort to make such an exact definition in a usable form began long ago, and discussions of this subject in textbooks have frequently referred to the nineteenth-century French mathematician Duhamel, who stated a theorem, one purpose of which was to justify (6) under certain conditions. The modern approach to the problem uses different language and is somewhat different in conception from the old form of Duhamel's theorem. We call the following theorem *Duhamel's principle*, because it is historically and pedagogically descended from the work of Duhamel.

THE PRINCIPLE OF DUHAMEL. *The limit relation* (6) *will be correct if the quantity $\Delta Q_i$ associated with the interval from $x_{i-1}$ to $x_i$ is such that the greatest of the expressions*

$$\left|\frac{\Delta Q_i}{\Delta x_i} - F(x_i')\right|$$

*approaches zero as $\|\Delta\| \to 0$.*

This form of the theorem is essentially due to the late Professor W. F. Osgood. We shall call it *Osgood's form of Duhamel's principle.*

In many situations where something like Duhamel's principle is needed, the formula of Bliss will meet the need. This formula can be deduced as a corollary of Osgood's form of Duhamel's principle.

**Example 2:** As an illustration of a situation where the formula of Bliss is not applicable, but Duhamel's principle is needed, let us consider the derivation of formula (7) in § 11-1, for arc length of a curve represented parametrically.

Going back to (3) in § 11-1, we employ the law of the mean on $x = \phi(t)$ and $y = \psi(t)$. If $(x_k, y_k)$ corresponds to $t_k$, and if $t_k - t_{k-1} = \Delta t_k$, then

$$x_k - x_{k-1} = \phi'(u_k)\,\Delta t_k, \qquad y_k - y_{k-1} = \psi'(v_k)\,\Delta t_k,$$

where $u_k$ and $v_k$ are certain numbers between $t_{k-1}$ and $t_k$, as provided by the law of the mean. Hence in this case

$$\overline{P_{k-1}P_k} = \{[\phi'(u_k)]^2 + [\psi'(v_k)]^2\}^{1/2}\,\Delta t_k.$$

In applying Duhamel's principle here, let $\Delta Q_k = \overline{P_{k-1}P_k}$. Then if $\Delta$ is the partition determined by $a = t_0 < t_1 < \cdots < t_n = b$, the definition of $L$ is

$$L = \lim_{\|\Delta\| \to 0} (\Delta Q_1 + \cdots + \Delta Q_n).$$

Now if $u_k$ and $v_k$ were the *same* point, we could take

$$F(t) = \{[\phi'(t)]^2 + [\psi'(t)]^2\}^{1/2} \tag{7}$$

and we would then have exactly $\Delta Q_k = F(u_k)\,\Delta t_k$. But since $u_k$ and $v_k$ may be different, the matter is not so simple. But since $\phi'$ and $\psi'$ are continuous, and since $u_k$ and $v_k$ get closer and closer together as $\|\Delta\| \to 0$, it can be shown that the conditions of Duhamel's principle are fulfilled, with $F$ given by (7) and

$$\text{maximum of } \left|\frac{\Delta Q_i}{\Delta t_i} - F(u_i)\right| \to 0 \qquad \text{as } \|\Delta\| \to 0.$$

This justifies the formula (7) in § 11-1.

## 11-4 The Area of a Surface of Revolution

As in § 6-1, let us consider a figure of revolution obtained by revolving a curve $y = f(x)$ about the $x$-axis. The curve generates what is called a *surface of revolution.* This surface forms the lateral boundary of a *solid* of revolution. We now set ourselves the problem of measuring the area of such a surface.

Let the interval $(a, b)$ of the $x$-axis be divided into $n$ parts $\Delta x_1, \cdots, \Delta x_n$, as usual; let ordinates $y_i = f(x_i)$ be erected at the division points $x_i$. Then draw the broken line joining the points in which these ordinates meet the curve (see Fig. 11-7). When the arc $AB$ is revolved about the $x$-axis, one of the chords composing the above mentioned broken line generates the lateral area of a frustum of a cone. Geometrical intuition suggests that, when the segments $\Delta x_i$ are small, the area of such a frustum of a cone

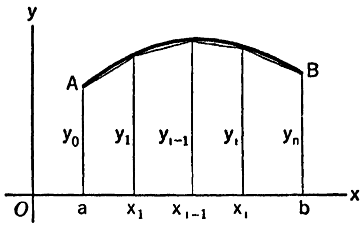



**Fig. 11-7**

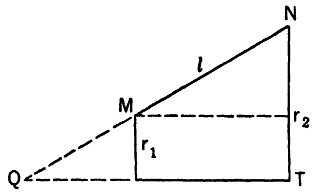



**Fig. 11-8**

is very nearly what we should mean by the area of the corresponding circular band on the surface itself. On the basis of this suggestion we *define* the area of the surface as the limit of the sum of the areas of all the conical frusta generated by the broken line.

Now, the lateral area of a frustum of a right circular cone is

$$S = \pi(r_1 + r_2)l, \tag{1}$$

where $r_1$, $r_2$ are the radii of the two bases, and $l$ is the slant height (see Fig. 11-8).* It follows from (1) that the area generated by the chord corresponding to the segment $\Delta x_i$ in Fig. 11-7 has the value

$$\pi(y_{i-1} + y_i)(\Delta x_i^2 + \Delta y_i^2)^{1/2} = \pi(y_{i-1} + y_i)\left[1 + \left(\frac{\Delta y_i}{\Delta x_i}\right)^2\right]^{1/2} \Delta x_i. \tag{2}$$

But $\Delta y_i = f(x_i) - f(x_{i-1}) = f'(u_i)\,\Delta x_i$, where $u_i$ is some point of the interval $(x_{i-1}, x_i)$, by the law of the mean (§ 2-1). Hence, expression (2) can be written in the form

$$\pi[f(x_{i-1}) + f(x_i)][1 + f'(u_i)^2]^{1/2}\,\Delta x_i. \tag{3}$$

The area $S$ of the surface of revolution will be the limit of the sum of the

* For in Fig. 11-8 the lateral area of the cone generated by revolving $QN$ about $QT$ is $\pi r_2 \cdot \overline{QN}$, while the area generated by $QM$ is $\pi r_1 \cdot \overline{QM}$. Hence, the area of the frustum generated by $MN$ is $S = \pi(r_2\overline{QN} - r_1\overline{QM})$. Let us now write $\overline{QN} = \overline{QM} + l$. Then,

$$S = \pi(r_2\overline{QM} + r_2 l - r_1\overline{QM}) = \pi(r_2 - r_1)\overline{QM} + \pi r_2 l.$$

Next, $(r_2 - r_1)\overline{QM} = r_1 l$, for by similar triangles,

$$\frac{r_2 - r_1}{l} = \frac{r_1}{\overline{QM}}.$$

Thus, finally, we obtain formula (1).

expressions (3) formed for $i = 1, 2, \cdots, n$. We can write this sum as two separate sums:

$$\pi\{f(x_0)[1 + f'(u_1)^2]^{1/2}\,\Delta x_1 + \cdots + f(x_{n-1})[1 + f'(u_n)^2]^{1/2}\,\Delta x_n\}$$
$$+ \pi\{f(x_1)[1 + f'(u_1)^2]^{1/2}\,\Delta x_1 + \cdots + f(x_n)[1 + f'(u_n)^2]^{1/2}\,\Delta x_n\}.$$

Each of these is a sum to which the formula of Bliss (§ 11-3) applies. They each, therefore, have the same limit, namely the definite integral

$$\pi \int_a^b f(x)[1 + f'(x)^2]^{1/2}\,dx.$$

Thus, finally, we see that the area of the surface of revolution is twice this integral, or

$$S = 2\pi \int_a^b y\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} dx, \tag{4}$$

where $y = f(x)$ is the equation of the generating curve.

The student will see that the surface area formula can be thought of in the form

$$S = 2\pi \int y\,ds$$

where $ds$ is to be calculated in terms of a convenient independent variable and its differential, and appropriate limits are to be supplied.

**Example:** Let a sphere be inside a circular cylinder whose radius is the same as that of the sphere. If two planes cut the cylinder at right angles to its axis, and intersect the sphere, show that the area on the sphere between the planes is the same as the area on the cylinder between the planes.

Think of the sphere as being generated by revolving the circle $x^2 + y^2 = a^2$ about the $x$-axis. Let the planes cut the $x$-axis at $x = c$ and $x = c + h$, where $h > 0$. To calculate $ds$ we have $2x\,dx + 2y\,dy = 0$, so

$$ds^2 = dx^2 + \left(\frac{-x}{y}\right)^2 dx^2 = \frac{y^2 + x^2}{y^2}\,dx^2 = \frac{a^2}{y^2}\,dx^2.$$

Then $2\pi y\,ds = 2\pi a\,dx$, and the required surface area on the sphere is

$$S = 2\pi \int_c^{c+h} a\,dx = 2\pi a h.$$

This answer, when interpreted, justifies the assertion about the areas on the sphere and the cylinder.

## EXERCISES

1. Find the surface area generated when the indicated arc is revolved about the $x$-axis.
   (a) $y = 2\sqrt{x}$ from $x = 0$ to $x = 8$;
   (b) $y = \frac{1}{3}x^3$ from $x = 0$ to $x = 2$;
   (c) $x = \sqrt{2}t^2$, $y = 2t$, from $t = 0$ to $t = 2$;
   (d) $y = \sin x$ from $x = 0$ to $x = \pi$;
   (e) $y = \frac{1}{2}(e^x + e^{-x})$ from $x = -1$ to $x = 1$.

**2.** Find the surface area generated when the indicated arc is revolved about the $y$-axis.

(a) $y = \log x$, from $x = 1$ to $x = 2\sqrt{2}$;

(b) $x = \cos 2y$, from $y = 0$ to $y = \frac{\pi}{4}$;

(c) $x^2 + y^2 = 25$, from $(4, 3)$ to $(3, 4)$.

**3.** Find the lateral surface of the cone generated by revolving the line $y = mx$ from $x = 0$ to $x = 1$ about the $x$-axis.

**4.** Find the area of the surface generated by revolving the arc of the cubical parabola $y = x^3$ from $(0, 0)$ to $(1, 1)$ about (a) the $x$-axis; (b) the $y$-axis.

**5.** Find the area generated by revolving the arch of the curve $y = \cos x$ from $x = -\pi/2$ to $x = \pi/2$ about the $x$-axis.

**6.** The arc of the parabola $y^2 = 4px$ from $(0, 0)$ to $(p, 2p)$ is revolved (a) about the $x$-axis; (b) about the $y$-axis. Find the area generated in each case.

**7.** Find the surface area generated by revolving the ellipse $x = a \cos t$, $y = b \sin t$ about (a) the $x$-axis; (b) the $y$-axis. The eccentricity of the ellipse is $e = \sqrt{a^2 - b^2}/a$.

**8.** One arch of the cycloid is revolved about its base. Find the area of the surface thus generated.

## 11-5 Moments of Mass Distributions. Center of Mass

Consider a system of $n$ particles, of masses $m_1, \cdots, m_n$, distributed in any fashion along a straight line, which we take to be the $x$-axis. Let the coordinate of $m_k$ be $x_k$. The product $m_k x_k$ is called the *moment*, or first moment, of $m_k$ relative to the origin (or about the origin). This moment is positive or negative according as $x_k > 0$ or $x_k < 0$, and it is zero if $x_k = 0$. The algebraic sum of all the moments is called the *total moment* of the system relative to 0. This total moment is

$$m_1x_1 + \cdots + m_nx_n. \tag{1}$$

Now let $M$ be the sum of all the masses. At what point $x = \bar{x}$ should a particle of mass $M$ be placed so as to have its moment relative to 0 the same as the total moment of the system? Evidently, if this is to be the situation, $\bar{x}$ is determined by the equation

$$M\bar{x} = m_1x_1 + \cdots + m_nx_n, \tag{2}$$

or

$$\bar{x} = \frac{m_1x_1 + \cdots + m_nx_n}{m_1 + \cdots + m_n}. \tag{3}$$

The point $x = \bar{x}$ is called the *center of mass* of the system.

The point $x = \bar{x}$ is also called the center of gravity of the system of particles, for the following reason. Suppose the $x$-axis is horizontal, and think of the segment of it which carries the masses as a light rod (so light as to have negligible weight in comparison with the total weight of the system of masses). Then if the rod is supported by suspending it on a cable attached at $x = \bar{x}$, the rod will balance in the horizontal position. This is because the tendency of the weights on the right of $x = \bar{x}$ to force that end of the rod down is exactly balanced by the tendency of the weights on the left of $x = \bar{x}$ to force the left end down. The algebraic sum of the moments about the point $x = \bar{x}$ is zero.

If a system of masses is distributed in the $xy$-plane, with mass particle $m_k$ at $(x_k, y_k)$, we define moments relative to the axes. In this case the sum (1) is called the total moment of the system about the $y$-axis (because $x_k$ is the algebraic distance from the $y$-axis to $m_k$). Likewise, the sum

$$m_1y_1 + \cdots + m_ny_n$$

is called the total moment of the system about the $x$-axis. In this case, if $M$ is the total mass, and if $(\bar{x}, \bar{y})$ is the point where a particle of mass $M$ must be placed in order for its moments $M\bar{x}$ and $M\bar{y}$, about the $y$-axis and $x$-axis, respectively, to be the same as the corresponding total moments of the system, this point $(\bar{x}, \bar{y})$ is called the center of mass of the system. Here $\bar{x}$ is given by (3) and $\bar{y}$ is given by an exactly similar formula with $y_k$ in place of $x_k$. The center of mass is also called the center of gravity in this case.

The center of mass concept can evidently be defined for systems of particles distributed in three-dimensional space. We need not spell out the details.

The center of mass concept is also used in other ways. What is meant by the phrase "the center of population of the United States"? For simplicity think of the land as a plane surface, and consider each person as a particle on the plane, at the location of his home. If each person in the United States is counted as a particle of unit mass, the center of population is the center of mass of this system. Of course, in practice, this center of population must be computed approximately, by using census data in appropriate lumps.

So much for distributions of a finite number of discrete particles. What about nondiscrete distributions? What is meant by the center of mass of a solid hemisphere, of a cone made of sheet metal, or of a coil spring? For purposes of mathematical study, we regard such objects as being composed of mass continuously distributed, either throughout a portion of space, or over a certain surface, or along a certain curve. Instead of mass particles, we have a *mass density* at each point, and the total mass, instead of being found as a finite sum, is found by calculating some kind of an integral of this density.

We shall presently learn how to find the centers of mass of certain continuous bodies by means of definite integrals. The general principle is the same in all cases, though the analytical details vary with the nature of the body. We are now concerned with the principle, not with the details.

Throughout the whole subject of mechanics, the gap between the notion of a system of particles and the notion of a continuous distribution of mass is bridged by the physical assumption that a continuous body may be treated as a limiting case of a finite system of particles. We give this assumption the following explicit form and adopt it as a governing principle:

A concept or physical law relating to a finite system of particles is to be carried over to the case of a continuous body by dividing the body into a number of pieces and imagining the mass of each piece to be concentrated as a particle at some one point of the piece. The resulting finite system of particles we shall refer to as *an auxiliary system of particles.* If now we consider the concept or physical law as it applies to the auxiliary system, the concept or physical law shall be carried over to the continuous body by increasing indefinitely the number of pieces in the auxiliary system in such a way that the maximum diameter of the pieces approaches zero.

As a particular case of the application of the governing principle, the center of mass of a continuous body is *defined* as the limiting position of the center of mass of the auxiliary system when the number of pieces increases indefinitely and their maximum diameter approaches zero. This definition of the center of mass of a body leads to the use of integrals for the calculation of the coordinates of the center of mass.

Finding the center of mass of a body is in many cases simplified by the use of the following theorem:

THEOREM 11-A. *If a body of mass $M$ consists of $n$ distinct parts, of masses $\Delta M_1, \cdots, \Delta M_n$, and if an auxiliary system of $n$ particles is formed by concentrating the mass of each of the parts at its own center of mass, then the center of mass of the entire body coincides with the center of mass of the auxiliary system.*

The content of the theorem is exactly expressed by the formula

$$M\bar{x} = \Delta M_1\bar{x}_1 + \cdots + \Delta M_n\bar{x}_n \tag{4}$$

and two similar formulas for $\bar{y}$ and $\bar{z}$. Here $\bar{x}$ is the abscissa of the center of mass of $M$, $\bar{x}_1$ is the abscissa of the center of mass of $\Delta M_1$, and so forth. The proof of the theorem, for solid bodies, is an immediate consequence of the formulas for $M$, $\bar{x}$, $\bar{y}$, $\bar{z}$ in terms of triple integrals, as given in Chapter XX.

When a body is composed of material whose density is the same throughout, the body is said to be *homogeneous.* The center of mass of a homogeneous body does not depend on the density, but only on the size and shape of the body; that is, upon its geometrical configuration. By the *centroid* of a geometrical configuration we mean the center of mass of the configuration when it is regarded as a homogeneous body.

EXERCISES

1. Locate the center of mass of masses: 10 at $(-\frac{1}{2}, 1)$, 4 at $(1, 0)$, 2 at $(2, -3)$ 8 at $(3, 2)$.
2. The position of the center of mass is not dependent upon the location of the axes. That is, if the axes are changed, the coordinates of the center of mass may change, but the point itself will not change. Show this (a) for translations of axes; (b) for rotations of axes. Refer to § 7-6.
3. Four equal masses are placed at the vertices of a parallelogram. Show that the center of mass is at the intersection of the diagonals.
4. Where is the center of mass of a system of three equal masses, placed one at each vertex of a triangle?

## 11-6 The Centroid of a Solid of Revolution

It is rather easily seen that when a body has an axis of symmetry its center of mass is on that axis. Hence the centroid of a solid of revolution is on the axis of the solid. We shall show how to find its position on the axis by the use of integration.

For definiteness, let the axis of revolution be the $x$-axis. We shall follow the notation of § 6-1, where we discussed the volume of solids of revolution. Let the solid be cut into $n$ thin circular disks by slicing it at right angles to the $x$-axis. If the constant density of the solid is $\rho$, the mass $\Delta M_i$ of the $i$th disk is $\Delta M_i = \rho \, \Delta V_i$, where $\Delta V_i$ is the volume of the disk. The centroid of $\Delta M_i$ is at a point $x = \bar{x}_i$ somewhere between the faces of the disk and on the $x$-axis. We now use Theorem 11-A of § 11-5 to give us the abscissa of the centroid of the entire solid:

$$\rho V \bar{x} = \rho \, \Delta V_1 \, \bar{x}_1 + \cdots + \rho \, \Delta V_n \, \bar{x}_n.$$

Hence, canceling $\rho$,

$$V\bar{x} = \Delta V_1 \, \bar{x}_1 + \cdots + \Delta V_n \, \bar{x}_n. \tag{1}$$

Let $y = f(x)$ denote the equation of the curve which, when revolved about the $x$-axis, generates the curved surface of our solid. The discussion in § 6-1 shows us that

$$\Delta V_i = \pi y_i'^2 \Delta x_i = \pi [f(x_i')]^2 \, \Delta x_i, \tag{2}$$

where $x_i'$ is some value of $x$ in the interval $\Delta x_i$. Thus

$$V\bar{x} = \pi(\bar{x}_1 y_1'^2 \, \Delta x_1 + \cdots + \bar{x}_n y_n'^2 \, \Delta x_n). \tag{3}$$

If we allow $n$ to increase indefinitely and the maximum $\Delta x_i$ to approach zero, the left side of (3) is unchanged, but the right side approaches a definite integral as its limit. In this way we see that

$$V\bar{x} = \pi \int_a^b xy^2 \, dx. \tag{4}$$

In the integration $y$ must be expressed in terms of $x$ from the equation of the curve. That the limit of the sum in (3) is the integral in (4) is assured by the formula of Bliss, § 11-3.

In solving problems the student should not merely depend upon formula (4). The general procedure is founded upon the cutting of the solid into pieces and the use of a formula such as (1). The procedure may of course be applied to the case of a solid of revolution about the $y$-axis, or any line. The volume elements $\Delta V$ must be expressed in appropriate coordinates. The form of the integral will then suggest itself immediately.

**Example 1:** The circle $x^2 + y^2 = a^2$ is revolved about the $x$-axis, generating a sphere. Find the centroid of the solid hemisphere for which $x \geqq 0$.

When Fig. 11-9 is revolved about the $x$-axis, the shaded strip generates a volume element. The student should see that this element has volume $\pi y^2 \Delta x$, where $y$ corresponds to a suitable value of $x$ between the two faces of the element. The centroid of this element is on the $x$-axis and between the faces of the element. The totality of such elements carries us from $x = 0$ to $x = a$. Hence, by the argument explained above,

$$V\bar{x} = \pi \int_0^a xy^2\,dx. \tag{5}$$

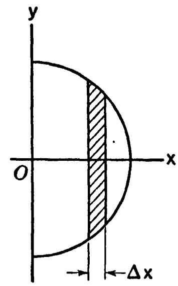



**Fig. 11-9**

We know that $V = \frac{2}{3}\pi a^3$. Also, since $y^2 = a^2 - x^2$,

$$\int_0^a xy^2\,dx = \int_0^a (a^2x - x^3)\,dx = \frac{a^4}{2} - \frac{a^4}{4} = \frac{a^4}{4}.$$

Thus, from (4), $\quad \dfrac{2}{3}\pi a^3\bar{x} = \pi\dfrac{a^4}{4}, \quad$ or $\quad \bar{x} = \dfrac{3}{8}a.$

Sometimes we find the volume of a solid of revolution by dividing it into cylindrical shells (cf. § 11-2). In such cases the centroid may also be calculated by the shell method of integration. The centroid of a cylindrical shell is on its axis, midway between the ends of the shell.

**Example 2:** If a spheroid is generated by revolving the ellipse $9x^2 + 16y^2 = 144$ about the $y$-axis, use the shell method to find the centroid of that half of the solid for which $y \geqq 0$.

A strip of width $\Delta x$ parallel to the $y$-axis generates a cylindrical shell. The volume of the shell is approximately $2\pi xy\,\Delta x$, and its centroid is on the $y$-axis,

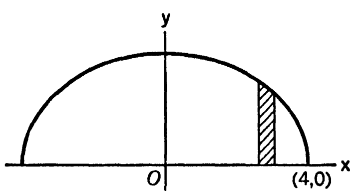



**Fig. 11-10**

a distance approximately $y/2$ above $O$, where $x$ is the arithmetic mean of the inner and outer radii of the shell and $y$ is the corresponding ordinate of the ellipse. Formula (1) is now replaced by

$$V\bar{y} = \bar{y}_1\,\Delta V_1 + \cdots + \bar{y}_n\,\Delta V_n,$$

where $\bar{y}_i$ is approximately $y_i/2$ and $\Delta V_i$ is approximately $2\pi x_i y_i\,\Delta x_i$. The typical term of the sum is thus approximately $\pi x_i y_i^2\,\Delta x_i$, and passage to the limit gives

$$V\bar{y} = \pi \int_0^4 xy^2\,dx = \pi \int_0^4 x\left(\frac{144 - 9x^2}{16}\right)dx.$$

We leave it for the student to verify that $V = 32\pi$, and to complete the integration and find $\bar{y} = \frac{9}{8}$.

## EXERCISES

Find the centroids of the solids of revolution in Exercises 1-9, first by the disk method and then by the shell method.

**1.** A right circular cone of altitude $h$ and radius of base $r$, with axis on the positive $x$-axis and vertex at the origin.

**2.** A right circular cone of altitude $h$ and radius of base $r$, with axis on the positive $y$-axis and center of the base at the origin.

**3.** The spherical segment cut from the sphere of illustrative Example 1 by planes $x = a - h$, $x = a$, where $0 < h \leqq 2a$. How can you check the answer?

**4.** The solid obtained by rotating that portion of the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ for which $x > 0$ about the $x$-axis.

**5.** The upper half of the solid generated when the ellipse of Exercise 4 is revolved about the $y$-axis.

**6.** The solid formed when the area between the parabola $x^2 = 4y$ and the $x$-axis, from $x = 0$ to $x = 4$, is revolved about the $x$-axis.

**7.** The solid formed when the area of Exercise 6 is revolved about the $y$-axis.

**8.** The solid formed when the area bounded by the parabola $y = x^2$ and the line $3y = 4x$ is revolved about the $x$-axis.

**9.** The solid formed by revolving the area of Exercise 8 about the $y$-axis. Use washer-shaped elements perpendicular to the $y$-axis.

**10.** In each part of this exercise a triangular area is specified by giving the vertices, and an axis is named. Locate the centroid of the solid formed by revolving the area about the axis. Do each part in two ways: once integrating with respect to $x$, and once with respect to $y$.

(a) $(0, 0)$, $(4, 0)$, $(4, 6)$; axis $x = 0$;
(b) $(0, 1)$, $(2, 3)$, $(0, 3)$; axis $y = 0$;
(c) $(-2, 1)$, $(2, 1)$, $(-2, 5)$; axis $y = 0$;
(d) $(0, -4)$, $(4, 4)$, $(0, 0)$; axis $x = 0$;
(e) the same triangle as in (d); axis $x = 4$;

(f) the same triangle as in (d); axis $y = 4$.

In (d), (e), (f) the $y$-integration calls for dividing the area into two parts by the line $y = 0$.

**11.** Consider the area between the parabola $y = 4x - x^2$ and the line $y = x$. Find the centroid of the solid generated when this area is revolved about (a) the axis $x = 0$; (b) the axis $y = 0$; (c) the axis $y = 4$; (d) the axis $x = 3$.

**12.** Consider the area between the curve $y = x^3$ and the $x$-axis, from $x = 0$ to $x = 1$. Find the centroid of the solid generated when this area is revolved about (a) the axis $x = 1$; (b) the axis $y = 1$; (c) the axis $x = 2$; (d) the axis $y = -1$.

**13.** Consider the area in the first quadrant bounded by the hyperbola $x^2 - y^2 = a^2$ and the line $x = 2a$. Find the centroid of the volume generated when this area is revolved about (a) the $x$-axis; (b) the $y$-axis; (c) the line $x = 2a$.

**14.** Find the centroid of the solid generated when the area cut from the first quadrant by the circle $x^2 + y^2 = a^2$ is revolved about the line $y = a$.

## 11-7 The Centroid of a Plane Area

Consider a thin sheet of material, such as a piece of paper, the bottom of a pie tin, or a flat strip of copper. For many purposes it is convenient and useful to regard such distributions of matter as being two-dimensional. In this section we shall deal with laminas. By a *lamina* is meant a plane area thought of as a two-dimensional spread of matter. For the present we shall deal with homogeneous laminas, i.e. those for which the mass of any part is directly proportional to the area of that part. The mass per unit area is called the density. Laminas of variable density are considered in Chapter XX.

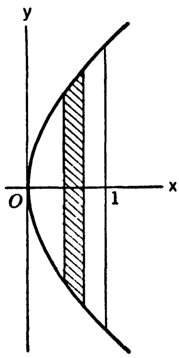


**Fig. 11-11**

To locate the centroid of a lamina, divide it into narrow strips in the manner described in § 6-5, that is, in the same way as when finding a plane area by integration. Then form an auxiliary system of particles by concentrating the mass of each strip at its center of gravity. Since the strip is approximately a long narrow rectangle, its centroid is approximately midway between the sides and halfway from one end of the strip to the other. The centroid of the auxiliary system is then found by Theorem 11-A of § 11-5, and a passage to the limit gives the coordinates of the centroid of the lamina in terms of integrals. The process is illustrated in the following examples.

**Example 1:** Find the centroid of the plane area bounded by the parabola $y^2 = 4x$ and the line $x = 1$.

From symmetry it is obvious that the centroid lies on the $x$-axis, so that $\bar{y} = 0$. To find $\bar{x}$, consider strips parallel to the $y$-axis. A typical strip of width $\Delta x_i$ and area $\Delta A_i$ is shown in Fig. 11-11. If the sides of the strip are the lines $x = x_{i-1}$ and $x = x_i$, it is clear that the centroid of the strip is at $x = \bar{x}_i$, where $\bar{x}_i$ is between $x_{i-1}$ and $x_i$. Also, by considering the equation of the parabola we see that the area of the strip is $4\sqrt{x_i'}\,\Delta x_i$, where $x_i'$ is between $x_{i-1}$ and $x_i$. By applying Theorem 11-A, § 11-5, to the auxiliary system formed with the strips, we find

$$A\bar{x} = \Delta A_1\,\bar{x}_1 + \cdots + \Delta A_n\,\bar{x}_n. \tag{1}$$

A typical term on the right has the structure

$$\Delta A_i\bar{x}_i = 4\bar{x}_i\sqrt{x_i'}\,\Delta x_i.$$

Hence, by the formula of Bliss, § 11-3, the limit of the sum in (1) is the integral

$$4\int_0^1 x\sqrt{x}\,dx.$$

Thus

$$A\bar{x} = 4\int_0^1 x^{3/2}\,dx = \frac{8}{5}.$$

The area $A$ itself is found by integration to be $\frac{8}{3}$. Hence $\frac{8}{3}\bar{x} = \frac{8}{5}$, or $\bar{x} = \frac{3}{5}$.

**Example 2:** Find the centroid of the area in Example 1 by integration with respect to $y$.

A typical strip of width $\Delta y_i$ is shown in Fig. 11-12. The length of the strip and its area can be expressed in terms of the distance from the $x$-axis to the strip. The distance from the $y$-axis to a point $(x, y)$ on the parabola $y^2 = 4x$ is $x = \frac{1}{4}y^2$. Hence, if the sides of the strip are $y = y_{i-1}$ and $y = y_i$, the length of the strip is approximately $1 - \frac{1}{4}y_i^2$ and its area is approximately $(1 - \frac{1}{4}y_i^2)\,\Delta y_i$. The exact area differs from this only by replacing $y_i$ by some value $y_i'$ between $y_{i-1}$ and $y_i$. The centroid of the strip is at a point midway between its ends, so that the abscissa of the centroid is approximately

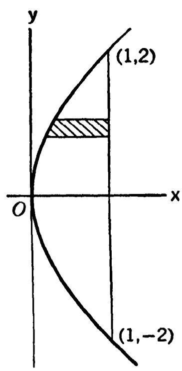


Fig. 11-12

$$\bar{x}_i = \frac{1}{2}(1 + x_i) = \frac{1}{2}\left(1 + \frac{1}{4}y_i^2\right).$$

Thus, approximately

$$\Delta A_i\,\bar{x}_i = \frac{1}{2}\left(1 - \frac{1}{4}y_i^2\right)\left(1 + \frac{1}{4}y_i^2\right)\Delta y_i = \frac{1}{2}\left(1 - \frac{1}{16}y_i^4\right)\Delta y_i.$$

Forming an auxiliary system in the usual way and passing to the limit, we have

$$A\bar{x} = \lim\,(\Delta A_1\,\bar{x}_1 + \cdots + \Delta A_n\,\bar{x}_n),$$

$$A\bar{x} = \frac{1}{2}\int_{-2}^{2}\left(1 - \frac{1}{16}y^4\right)dy.$$

Calculation of this integral leads to the result $\bar{x} = \frac{3}{5}$ as before.

It is evident that the methods here illustrated may be used to find $\bar{y}$ for a plane area.

### EXERCISES

1. Consider the first quadrant half of the area occurring in the illustrative examples. Find $\bar{x}$ and $\bar{y}$ for this area, using two different methods in each case.

In Exercises 2-8 find the centroid of each area, using two methods for each coordinate. Take advantage of symmetry wherever possible, and do not compute the areas by integration if they are known from standard formulas.

2. The right triangle formed by the lines $y = x$, $x = 1$, $y = 0$.
3. The area bounded by the parabola $3y^2 = -4(x - 3)$ and the $y$-axis.
4. The area bounded by the parabola $Hx^2 = -B^2(y - H)$ and the $x$-axis ($B$ and $H$ positive).
5. The right half of the circular area bounded by $x^2 + y^2 = a^2$.
6. The area in the first quadrant bounded by the ellipse $16x^2 + 9y^2 = 144$.
7. The triangle with vertices at $(0, 0)$, $(a, 0)$, and $(b, c)$, where $a$, $b$ and $c$ are positive.
8. The smaller area bounded by the circle $x^2 + y^2 = 25$ and the line $x + y = 5$.
9. Find the centroids of each of the following areas.
   (a) Between $y = 6x - x^2$ and $y = x$;
   (b) between $x = 4y - y^2$ and $y = x$;
   (c) between $y = x^2$ and $y - x = 2$;
   (d) between $y = x^3 - 3x$ and $y = x$ and on the right of the $y$-axis,
   (e) the trapezoid bounded by $x - 2y + 8 = 0$, $x + 3y + 5 = 0$, $x = -2$, $x = 4$;
   (f) the trapezoid with vertices $(5, -1)$, $(8, -1)$, $(7, 6)$, $(-2, 6)$.
10. Find the centroid of the plane region defined by $0 \leq y \leq \sin x, 0 \leq x \leq \pi$.
11. Find the centroid of the area in the first quadrant and within the four-cusped hypocycloid $x^{2/3} + y^{2/3} = a^{2/3}$. Use an appropriate trigonometric substitution to calculate the integrals.
12. Find the centroid of the smaller area cut from the inside of the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ by the line $bx + ay = ab$.

## 11-8 Forces and Fluid Pressure

In this section we consider how to calculate the total force which is exerted by a fluid such as water on a given portion of a vertical wall which forms part of the container of the fluid. Our analysis would apply, for example, to the force exerted on one end of a swimming pool by the water in the pool, or to the force exerted on one end of a cylindrical tank lying on its side by gasoline partly filling the tank.

A horizontal surface submerged in a fluid at rest is subjected to a downward force equal in amount to the weight of the column of fluid directly above the surface. Thus, for example, the force exerted by water on a square foot of the bottom of a pool 8 feet deep amounts to the weight of 8 cubic feet of water, or $8(62.4) = 499.2$ pounds.

The force per unit area at depth $h$ in a fluid is $wh$, where $w$ is the weight per unit volume of the fluid. This force per unit area is called *fluid pressure*. It is a physical law that the pressure of a fluid at a point in it is exerted equally in all directions. This means for instance, that at the bottom of a side wall of a pool, the force exerted by the water on a square inch of the bottom is *practically* the same as the force exerted on a square inch of the side wall right at the bottom. But not *exactly* the same, because whereas the square inch of bottom is all at the lower depth, the depth varies over the square inch of side wall, and the pressure is slightly less one inch from the bottom than at the bottom.

We shall now attack our general problem. Consider a vertical plane surface with fluid pressing on one side of it. We wish to find the total force on a specified part of this vertical plane. Suppose this part is outlined as shown in Fig. 11-13, the depth varying from $x = a$ to $x = b$, and the width of the specified part being $u$ at depth $x$. Then $u$ will be some function of $x$ which we can compute if we know the equations of the curves which bound the part in question of the plane.

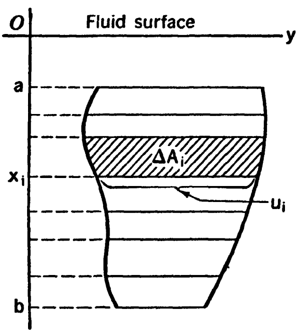



**Fig. 11-13**

We imagine narrow horizontal strips to be drawn across the surface. Let $\Delta A_i$ be the area of the strip between $x = x_{i-1}$ and $x = x_i$. From the physics of the situation we see that the force with which the fluid presses on this strip is more than $wx_{i-1}\,\Delta A_i$ but not so much as $wx_i\Delta A_i$. Hence this part of the total force $F$ is expressible as $\Delta F_i = wx_i'\,\Delta A_i$, where $x_{i-1} < x_i' < x_i$. Also, we can express the area of the strip in the form $\Delta A_i = u_i'\,\Delta x_i$, where $u_i'$ is the average value of $u$ as $x$ varies from $x_{i-1}$ to $x_i$. Hence the total force is

$$F = w(x_1'u_1'\,\Delta x_1 + \cdots + x_n'u_n'\,\Delta x_n).$$

On passing to the limit in the usual manner, and applying the formula of Bliss, we obtain

$$F = w\int_a^b xu\,dx. \tag{1}$$

It is not necessary to take the origin in surface of the fluid. If some other arrangement of axes is chosen, however, it must be remembered that the $x$ in (1) means the distance from the surface of the fluid down to a typical horizontal strip.

There is a relation between the force given by (1) and the position of the centroid of the submerged plane area on which we are computing the force. If $A$ is the number of square units of this area, and if the centroid is $\bar{x}$ units below the surface, we know from § 11-7 that

$$A\bar{x} = \int_a^b xu\,dx. \tag{2}$$

On comparing (1) and (2), we see that

$$F = wA\bar{x}. \tag{3}$$

The student should bear in mind the meaning of $A$ and $\bar{x}$ in this formula. If $A$ and the position of the centroid are already known, $F$ can be computed at once from (3). Otherwise, we work with the integral in (2).

**Example:** The trapezoidal area $ABCD$ shown in Fig. 11-14 is the end of a tank for storing water. If the tank is filled up to the line $y = 4$, find the total force with which the water presses on the end of the tank.

A typical strip of width $\Delta y$ is shown. To get the relation between $x$ and $y$ at the end of the strip we need the equation of the line $BC$. This equation is easily found to be $y = 3x - 9$. The area of the strip is approximately $x\,\Delta y$, and its distance below the water surface is $4 - y$. Hence the force on the strip is approximately $w(4 - y)x\,\Delta y$, and the total force is

$$F = w\int_{-3}^{4} (4 - y)x\,dy,$$

where $x = \dfrac{y + 9}{3}$. This works out to be

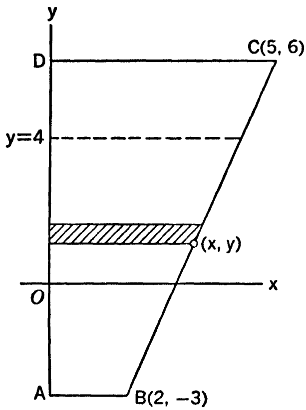


**Fig. 11-14**

$$F = \frac{1225w}{18} = \frac{1225}{18}(62.4) = 4246\frac{2}{3}\text{ lb}.$$

The details are left to the student.

## EXERCISES

**1.** Find the total force due to fluid pressure on one side of each of the following areas. In each case assume the $y$-axis is horizontal and the positive $x$-axis extends downward. The location of the fluid surface is specified in each case.

(a) Area bounded by $y^2 = x$, $x = 4$; fluid surface at $x = 0$.

(b) Same as (a), but with fluid surface at $x = -1$.
(c) Area bounded by $y^2 = -x$, $x = -9$, fluid surface at $x = -9$.
(d) Area bounded by $\sqrt{3}y = x - 1$, $\sqrt{3}y = -(x - 1)$, $x = 4$; fluid surface at $x = 0$.
(e) Same as (d), but with fluid surface at $x = 1$.
(f) Triangular area with vertices $(1, \pm 3)$, $(5, 0)$; fluid surface at $x = 0$; at $x = 1$.
(g) Trapezoidal area with vertices at $(2, \pm 3)$, $(5, \pm 1)$; fluid surface at $x = 0$; at $x = -1$.
(h) Area between $2y = x^2$ and $y = 8$, and below $x = 2$; fluid surface at $x = 2$.
(i) The upper and lower halves (separately) of the circular area bounded by $(x - 2)^2 + y^2 = 4$; fluid surface at $x = 0$.

2. In this exercise it is assumed that the $x$-axis is horizontal and the positive $y$-axis extends upward. Find the total force due to fluid pressure on one side of each of the following areas, with the location of the fluid surface as specified.
(a) Area bounded by $16y = x^2$, $y = 4$; fluid surface at $y = 4$.
(b) Area bounded by $x^2 + y^2 = 25$; fluid surface at $y = 10$. Solve by integration and check by use of the position of the centroid.
(c) Area bounded by $2x + 3y = 24$, $x = 0$, $y = 0$; fluid surface at $y = 12$.
(d) Area bounded by $2y^2 = 5x$ and $x = 10$; fluid surface at $y = 5$.

3. Find the force on the end of a swimming pool $b$ feet wide and $h$ feet deep.

4. A cylindrical tank 3 feet in diameter is lying on its side. Find the total force due to water pressure on one end of the tank, (a) if the tank is half full of water; (b) if the water is $\frac{3}{4}$ foot deep.

5. The outlet gate of a reservoir closes a circular hole in the side. Find the force on the outlet gate if the hole is 4 feet in diameter and its center is 40 feet below the water level.

6. A cylindrical tank 8 feet in diameter is lying on its side. If it contains water to a depth of 6 feet, find the total force due to the water pressure on one end of the tank.

7. A parabolic plate is lowered, vertex downward, until the latus rectum lies in the surface of a liquid. Find the force on one side of the plate, if the latus rectum is 4 feet long.

8. An elliptical plate, major axis 6 units, minor axis 4 units, is submerged vertically until the minor axis lies in the surface of the water. Find the force on one side of the submerged portion.

9. Find the force on one side of the submerged portion of the elliptical plate of Exercise 8 if its center is 2 units below the surface, and the major axis is horizontal.

10. If, in Exercise 9, the minor axis is horizontal and the center 3 units below the surface. find the force on one side of the submerged portion.

**11.** The rectangular endgate of a trough is to be mounted so that it may be turned about a horizontal axis in its own plane. Where should this axis of support be placed so that the gate will not tend to rotate when the trough is full of water?

**12.** Formulate a general method for solving problems like that of Exercise 11. *Hint:* Find an axis such that the algebraic sum of the moments about this axis, due to the pressure of the fluid on the various horizontal strips $\Delta A_i$ (Fig. 11-13) is zero.

## 11-9 More on Mass Distributions and Centroids

Consider the surface which is generated by revolving a curve $y = f(x)$ about the $x$-axis. We consider the part of the surface corresponding to $a \leq x \leq b$. It is assumed that the values of $f(x)$ are nonnegative and that $f$ has a continuous derivative. Just as we formed the idealized conception of a two-dimensional plane distribution of matter in § 11-7, so now we can imagine our surface of revolution to be a curved lamina, bearing mass continuously distributed over it in a thin layer. We shall assume the density to be constant, so that the lamina is what we call *uniform*, or homogeneous. Then the centroid of the surface lies on the $x$-axis, which is the axis of symmetry. If $S$ is the total area of the surface, the coordinate $\bar{x}$ of the centroid is given by

$$S\bar{x} = 2\pi \int_a^b xy \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx. \tag{1}$$

The reason for this will now be given. Consider Fig. 11-15, which shows a longitudinal section of our surface by a plane through the axis of symmetry, and a thin slice made by two planes perpendicular to this axis.

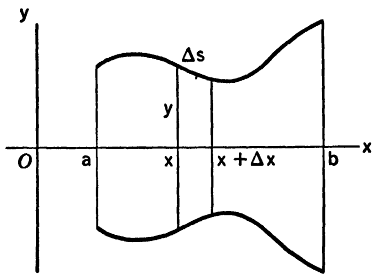



**Fig. 11-15**

This slice cuts a thin, ribbonlike circular band from the surface. This ribbon band is generated by revolving about the $x$-axis a segment of arc length $\Delta s$. The area $\Delta S$ of this ribbon band is approximately $\Delta S = 2\pi y\, \Delta s$, and its moment relative to the $y$-axis is approximately $k(2\pi y\, \Delta s)x$, where

$k$ is the constant mass per unit area. Consequently, by the usual limiting process, we expect the total moment to be

$$2\pi k \int_0^l xy\, ds,$$

where $s = 0$ when $x = a$ and $s = l$ when $x = b$. This total moment must equal $kS\bar{x}$, and so, canceling $k$, we obtain

$$S\bar{x} = 2\pi \int_0^l xy\, ds. \tag{2}$$

On changing the variable of integration from $s$ to $x$ and recalling the formula for $ds$, we obtain (1). The details of this argument can be filled in more precisely by using upper and lower estimates of the exact moment of each of the ribbon bands. In practice we may start from (2) and compute $ds$ in terms of any convenient parameter.

**Example 1:** The segment of the line $3y = x + 3$ from $x = 0$ to $x = 3$ is revolved about the $x$-axis, generating a frustum of a cone. Locate the centroid.

Here $3\, dy = dx$, and (1) becomes

$$S\bar{x} = 2\pi \int_0^3 x\left(\frac{x+3}{3}\right)\sqrt{1 + \frac{1}{9}}\, dx.$$

The area itself is

$$S = 2\pi \int_0^3 \frac{x+3}{3}\sqrt{1 + \frac{1}{9}}\, dx.$$

Calculating, we find $S\bar{x} = 5\pi\sqrt{10}$, $S = 3\pi\sqrt{10}$, $\bar{x} = \frac{5}{3}$. Details are left to the student.

*Thin Wires of Varying Density*

The subject of linear density of thin wires was discussed in § 6-10. This discussion should be reread at the present time. The notion of linear density applies to curved wires as well as to straight ones. If $\sigma$ is the linear density along a curve, the mass of the curve is

$$M = \int_0^l \sigma\, ds, \tag{3}$$

where $s$ is arc length measured from $s = 0$ at one end to $s = l$ at the other. If the curve is in the $xy$-plane, its center of mass $(\bar{x}, \bar{y})$ can be located by means of the formulas

$$M\bar{x} = \int_0^l x\sigma\, ds, \qquad M\bar{y} = \int_0^l y\sigma\, ds. \tag{4}$$

Actual integration may be carried out in terms of some parameter other than $s$, by changing variables in the standard way.

**Example 2:** Consider a material wire bent into the form of the parabolic arc $y = x^2$ from $x = -1$ to $x = 1$. Suppose the density is $\sigma = |x|$. Locate the center of mass.

Because of symmetry it is clear that $\bar{x} = 0$. We shall do the integrations with respect to $x$, and for this purpose

$$ds = \sqrt{1 + y'^2}\, dx = \sqrt{1 + 4x^2}\, dx.$$

Hence

$$M = \int_{-1}^{1} |x| \sqrt{1 + 4x^2}\, dx, \qquad M\bar{y} = \int_{-1}^{1} x^2|x| \sqrt{1 + 4x^2}\, dx.$$

Because the integrands are even functions,

$$M = 2\int_0^1 x\sqrt{1 + 4x^2}\, dx, \qquad M\bar{y} = 2\int_0^1 x^3\sqrt{1 + 4x^2}\, dx.$$

The integrals are easily calculated by the substitution $u = \sqrt{1 + 4x^2}$, $u^2 = 1 + 4x^2$, $u\, du = 4x\, dx$. We leave details to the student. The results are

$$M = \frac{1}{6}(5\sqrt{5} - 1), \qquad M\bar{y} = \frac{5\sqrt{5}}{12} + \frac{1}{60}.$$

This makes $\bar{y}$ slightly less than 0.56.

## EXERCISES

**1.** If the semicircle $x = \sqrt{a^2 - y^2}$ is revolved around the $x$-axis, find the centroid of the resulting hemispherical surface.

**2.** Prove that the centroid of the lateral surface of a right circular cone is on the axis, two thirds of the way from the vertex to the base. Work the problem in two ways:
(a) By placing the cone with vertex at the origin and axis along the $x$-axis.
(b) By placing the cone with the center of its base at the origin and axis along the $y$-axis.

**3.** When the ellipse $3x^2 + 4y^2 = 12$ is revolved about the $x$-axis, the half for which $x \geq 0$ generates a surface of area $S = \frac{\pi}{3}(9 + 2\pi\sqrt{3})$. Find the centroid of this surface.

**4.** Prove that the centroid of a zone of a spherical surface is on the axis of symmetry, halfway between the bases of the zone.

**5.** Find the center of mass of a homogeneous wire in the shape of the semicircular arc $y = \sqrt{a^2 - x^2}$. Use two methods:
(a) Integrating with respect to $x$.
(b) Integrating with respect to $\theta$, where $x = a\cos\theta$, $y = a\sin\theta$.

**6.** Locate the centroid of the arc of one arch of the cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$.

**7.** If the parabolic arc $y = x^2$ from $x = -1$ to $x = 1$ is a homogeneous wire, locate its centroid and compare your answer numerically with that of Example 2.

**8.** A wire has the shape of a semicircle of radius 2 feet. The density at a point of the wire varies in direct proportion to the distance from that point to the straight line joining the two ends of the wire. If the maximum density is $\frac{1}{4}$ pound per foot, find the mass of the wire and locate its center of mass.

**9.** Locate the center of mass of the first quadrant arc of the circle $x^2 + y^2 = a^2$ if it is a material wire of density $\sigma = x$.

**10.** Locate the centroid of the first quadrant arc of the four-cusped hypocycloid $x = a \cos^3 \theta$, $y = b \sin^3 \theta$.

*CHAPTER* XII

# POLAR COORDINATES

## 12-1 Elements of the Use of Polar Coordinates

A very simple way to introduce the subject of polar coordinates is the following. Consider any particular point $P$, except the origin, in the $xy$-plane. Let $r$ be the positive distance $\overline{OP}$, and let $\theta$ be any angle (in radian measure) such that

$$x = r\cos\theta, \qquad y = r\sin\theta, \tag{1}$$

where $(x, y)$ are the rectangular coordinates of $P$. See Fig. 12-1. There is not a unique determination of $\theta$. The meaning of equations (1) is that $\theta$ is the angle from the positive $x$-axis to the ray from $O$ through $P$. If $\theta_0$ is an acceptable measure of this angle, so is $\theta_0 \pm 2\pi$, $\theta_0 \pm 4\pi$, and so on, so that there are an infinite number of possibilities. The number $(r, \theta)$ (customarily written in this order) are called *a set of polar coordinates of P*. A little later on we shall see about the possibility of having negative values of $r$.

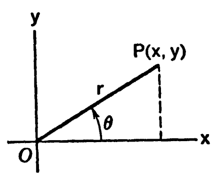


**Fig. 12-1**

**Example 1:** If $P$ has rectangular coordinates $(-\sqrt{3}, -1)$, we have $r = 2$. One possible choice of $\theta$ is $7\pi/6$. Another is $-5\pi/6$. (The student should draw a figure.) In general, for $\theta$ we may choose $7\pi/6 + 2\pi k$, where $k$ is any integer, positive, negative, or zero.

We placed the restriction that $P$ not be the origin. What if it is? Then $r = 0$, and in this case *any* choice of $\theta$ will satisfy equations (1). Hence for any $\theta$, $(0, \theta)$ is considered acceptable as a set of polar coordinates of the origin.

The fact that a point does not determine a unique set of polar coordinates is somewhat of a nuisance. It *is* true, however, that when any particular pair of polar coordinates of a point are given, we can use them to locate the point with certainty.

Our first aim is to become familiar with the use of polar coordinates in studying certain curves. There are two aspects of this sort of thing. We may take a given equation involving $r$ and $\theta$, and by interpreting $(r, \theta)$ as polar coordinates of a point, we may find the graph which consists of all points arising from pairs $(r, \theta)$ which satisfy the equation. Or we may start with a given curve (perhaps described by geometrical requirements, or perhaps defined as the locus of an equation in $x$ and $y$), and from this we may seek to find an equation involving $r$ and $\theta$ which must be satisfied by at least one set of polar coordinates of every point on the given curve. The relation between equations and graphs is not as simple with polar coordinates as with rectangular coordinates, because of the fact that a point has many sets of polar coordinates.

As we shall see in a moment, it is very natural to extend the concept of polar coordinates in such a way as to admit negative values of $r$. We explain this in the last part of our consideration of the next example.

**Example 2:** Consider the circle of radius $b$ with center at $x = b$, $y = 0$ (see Fig. 12-2).

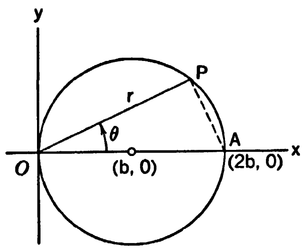



**Fig. 12-2**

The equation of the circle in rectangular coordinates is

$$x^2 + y^2 - 2bx = 0. \tag{2}$$

If $P$ is on the circle and $(r, \theta)$ are polar coordinates of $P$ (with $r \geq 0$) then, on substituting equations (1) into the equation (2), we obtain

$$r^2 \cos^2 \theta + r^2 \sin^2 \theta - 2br \cos \theta = 0, \quad \text{or} \quad r(r - 2b \cos \theta) = 0.$$

Hence either $r = 0$, or else

$$r = 2b \cos \theta. \tag{3}$$

This equation (3) is an equation for the circle in polar coordinates, in the following sense: if we let pairs $(r, \theta)$ be generated by selecting $\theta$ arbitrarily and computing $r$ from (3), all the pairs for which $r \geq 0$ lead to points on the circle, and every point on the circle is obtained from some pair. We observe that if we let $\theta$ increase from $-\pi/2$ to $\pi/2$, $r$ increases from 0 to $2b$ and then decreases back to 0. Hence, as $\theta$ goes from $-\pi/2$ to $\pi/2$, the point goes once around the circle in the counterclockwise sense. We know that all the points we get in this way *are* on the circle by the following argument: Suppose (3) holds. Put this $r$ in (1) and show that $x$ and $y$ satisfy (2). The details are

$$x = 2b\cos^2\theta, \qquad y = 2b\sin\theta\cos\theta,$$
$$x^2 + y^2 - 2bx = 4b^2\cos^4\theta + 4b^2\sin^2\theta\cos^2\theta - 4b^2\cos^2\theta$$
$$= 4b^2\cos^2\theta\,(\cos^2\theta + \sin^2\theta - 1) = 0.$$

Alternatively, we may make a geometric argument based on Fig. 12-2 and the fact that when $P$ is on the circle, angle $OPA$ is a right angle.

In practice, a great deal of the usefulness of polar coordinates comes from using them directly to express geometrical relations, and we shall not usually deal at all with the equation of a curve in rectangular coordinates if we can do all that is needed directly with an equation in polar coordinates.

Two things remain to be pointed out about the circle and equation (3). *First:* Some of the sets of polar coordinates of the origin do not satisfy equation (3). The only ones which do are those for which $\cos\theta = 0$.

*Second:* If $\theta$ is an angle of the second or third quadrant, (3) gives a negative value for $r$. In this case the pair $(r, \theta)$ is interpreted as a set of polar coordinates of the point $P$ located as follows: Draw the ray from 0 located by the angle $\theta$; extend it backward to form a complete straight line through $O$, and let $P$ be on this extension, a distance $-r$ from $O$. See Fig. 12-3. This point $P$ also has $(-r, \theta + \pi)$ as a set of polar coordinates, of course. It is on the circle.

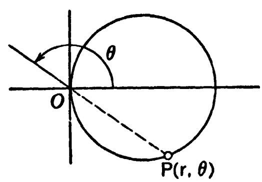


**Fig. 12-3**

The foregoing discussion, although based on the particular case of the circle and equation (3), illustrates the general idea of interpreting $(r, \theta)$ as a set of polar coordinates when $r < 0$. In our studies of plotting a curve from an equation in polar coordinates we shall always interpret negative $r$'s in this way.

Most of the curves which we consider in connection with polar coordinates are defined by equations whose general form is $r = f(\theta)$. The function $f$ is usually rather simple, and we can make a satisfactory sketch of the curve by examining the way in which $f(\theta)$ changes as $\theta$ varies. Certain points on the graph should be located by tabulating pairs $(r, \theta)$ which satisfy the equation. But an effort should be made to tabulate pairs $(r, \theta)$ which do the most possible in contributing to an effective visualization of the curve. The rest of the work should not be in plotting more points, but in discovering the essential characteristics of the function $f$.

For most simple graphs it is not necessary to use calculus in the construction of the graph, though calculus may be helpful in certain respects. Among the important matters are those asked about in the following list:

For what values of $\theta$ is $f$ defined?

When is $f(\theta)$ positive, when zero, and when negative?

For what values of $\theta$ does $f(\theta)$ reach maximum and minimum values?

Is there any kind of symmetry of the graph which can be discovered by examining the equation?

Probably the most important single consideration is that of knowing when $f(\theta)$ is increasing and when it is decreasing, as $\theta$ increases. This information can be worked out by studying $f'(\theta)$, if necessary.

**Example 3:** Sketch the curve $r = \sqrt{3} - \sin\theta$ (called a limaçon).

In this case we see that all values of $\theta$ are admissible, and that $r$ is always positive. Because of the periodicity of the sine, a study of what happens as $\theta$ goes from 0 to $2\pi$ will be adequate. Now the largest value of $r$ comes when $\sin\theta = -1$, i.e., when $\theta = 3\pi/2$, and the smallest value of $r$ comes when $\sin\theta = 1$, i.e., when $\theta = \pi/2$. We tabulate $(r, \theta)$ as shown, using $\theta =$ multiples of $\pi/2$.

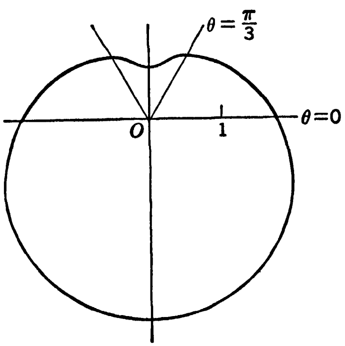



Fig. 12-4

| $\theta$ | $r$ |
|---|---|
| 0 | $\sqrt{3}$ |
| $\frac{\pi}{2}$ | $\sqrt{3} - 1$ |
| $\pi$ | $\sqrt{3}$ |
| $\frac{3\pi}{2}$ | $\sqrt{3} + 1$ |
| $2\pi$ | $\sqrt{3}$ |

Then we draw in the curve smoothly, noting that $r$ decreases as $\theta$ goes from 0 to $\pi/2$, increases as $\theta$ goes from $\pi/2$ to $3\pi/2$, and decreases again as $\theta$ goes from $3\pi/2$ to $2\pi$. The curve has one feature which is not readily detectable by the foregoing simple procedure. This is the "dimple" on the top of the curve. One way to discover something about this dimple, if its presence is suspected, is the following. *Consider how $y$ varies as $\theta$ varies.* Now

$$y = r\sin\theta = \sqrt{3}\sin\theta - \sin^2\theta,$$

$$\frac{dy}{d\theta} = \sqrt{3}\cos\theta - 2\sin\theta\cos\theta = \cos\theta\,(\sqrt{3} - 2\sin\theta).$$

From this we see when $y$ is increasing and when it is decreasing. We see that the critical values of $\theta$ are those for which $\cos\theta = 0$ or $\sin\theta = \sqrt{3}/2$.

These correspond to relative maximum or minimum values of $y$. Now, $\cos \theta = 0$ yields $\theta = \pi/2$, $3\pi/2$. The other critical values of $\theta$ are $\pi/3$ and $2\pi/3$. Note that $y$ decreases as $\theta$ goes from $\pi/3$ to $\pi/2$.

We also note the symmetry of the curve relative to the $y$-axis. We always have this kind of symmetry when $r$ is expressed as a function exclusively of $\sin \theta$.

We occasionally deal with curves whose equations have the form $r^2 = f(\theta)$. In this case, if $\theta$ is such that $f(\theta) < 0$, there is no corresponding point on the graph, since we must have $r^2 \geq 0$. But if $\theta$ is such that $f(\theta) > 0$, there are *two* corresponding points on the graph, with $r = \pm\sqrt{f(\theta)}$. These points are symmetrically placed relative to the origin. Hence a graph of $r^2 = f(\theta)$ is always symmetric with respect to $O$.

**Example 4:** Consider the curve $r^2 = a^2 \cos 2\theta$, $a > 0$ (called a lemniscate).

It suffices to consider the situation when $\theta$ goes from 0 to $\pi$, for when $\theta$ goes from $-\pi$ to 0, $\cos 2\theta$ does the same things as if $\theta$ were going from $\pi$ to 0, owing to the fact that $\cos(-2\theta) = \cos 2\theta$. The curve is therefore symmetric with respect to the $x$-axis. Such is always the case when $r$ is expressed entirely in terms of cosines of $\theta$ or multiples of $\theta$.

As $\theta$ increases from 0, $\cos 2\theta$ starts at 1 and decreases, reaching 0 when $2\theta = \pi/2$, or $\theta = \pi/4$. Between $\theta = \pi/4$ and $\theta = 3\pi/4$ we get no graph, because $\cos 2\theta < 0$. As $\theta$ goes from $3\pi/4$ to $\pi$, $\cos 2\theta$ increases from 0 to 1. The curve is shown in Fig. 12-5.

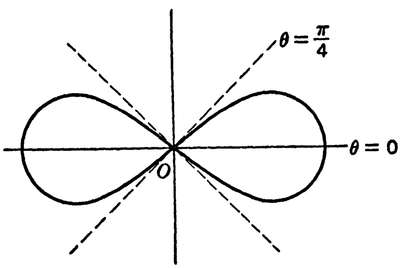



**Fig. 12-5**

We sometimes have to find intersections of curves which are defined by equations in polar coordinates. As a general problem this can be rather awkward, because it cannot always be completely solved by solving the two equations simultaneously. The reason for this is that a point can be on each of two curves and yet not have a pair of polar coordinates which satisfies both equations simultaneously. An extreme case of this is furnished by the two equations

$$r = 1 + \sin^2 \theta, \qquad r = -1 - \sin^2 \theta,$$

*which define the same curve.* Yet all of the $r$'s in one case are positive and all in the other case are negative. Another example:

$$r = 1 + \cos\theta, \qquad r^2 = 4\cos 2\theta.$$

Here the origin is on both curves, in the first case with coordinates $(0, n\pi)$, where $n$ is any odd integer, and in the second case with coordinates $(0, \pi/4 + k\pi/2)$, where $k$ is any integer.

What then is to be done about finding intersections? It is possible to develop general rules, but it is not worth while for what we need to do. In practice we shall rely on having good enough graphs of the curves to see whether there are any intersections. And when there are, we shall ordinarily be able to find the points of intersection either by solving simultaneous equations or by seeing where the points are directly from a figure.

## EXERCISES

**1.** Plot the following curves and explain how you know that each is a circle.
(a) $r = 8\sin\theta$. (c) $r = 3$.
(b) $r = -4\cos\theta$. (d) $r = -6\sin\theta$.

**2.** Draw and identify the graph in each case.
(a) $r = 4\csc\theta$. (c) $r = -2\csc\theta$.
(b) $r = 2\sec\theta$. (d) $r = -5\sec\theta$.

**3.** The following curves are called *cardioids*. Plot the first one carefully, and then plot the others, noting the way in which the form of the equation changes as the position of the curve is changed. The point of the curve at $O$ is called a cusp.
(a) $r = a(1 + \cos\theta)$. (c) $r = a(1 - \cos\theta)$.
(b) $r = a(1 + \sin\theta)$. (d) $r = a(1 - \sin\theta)$.

**4.** Plot the lemniscate $r^2 = b^2 \sin 2\theta$.

**5.** The curves $r = a + b\cos\theta$ with $ab \neq 0$ are called limaçons. In the special case $|a| = |b|$ they are cardioids. If $a > b$, the curve has a general resemblance to that in Fig. 12-4, but there is not always a dimple. If $a < b$, the curve intersects itself, forming a loop inside the larger part of the curve. Plot the following limaçons.
(a) $r = 1 + 2\sin\theta$. (c) $r = 2\sqrt{2} + 2\cos\theta$.
(b) $r = 1 + \sqrt{2}\cos\theta$. (d) $r = 5 + 2\sin\theta$.

**6.** Curves with equations of the form $r = a\cos n\theta$ or $r = a\sin n\theta$ are called *roses*. If $n$ is an odd integer there are $n$ lobes. If $n$ is an even integer there are $2n$ lobes. Plot each curve.
(a) $r = a\cos 3\theta$. (c) $r = a\cos 2\theta$.
(b) $r = a\sin 2\theta$. (d) $r = a\sin 3\theta$.

**7.** Plot each curve.
(a) $r = 2 + \sin 2\theta$. (c) $r^2 = 4\cos\theta$.
(b) $r = 4\sin^2\theta$. (d) $r^2 = 4\sin\theta$.

**8.** Find the largest and smallest values of $x$ on the limaçon $r = 6 - 2\sqrt{2}\sin\theta$.

**9.** Find the largest and smallest values of $y$ on
(a) The cardioid $r = 2(1 - \cos\theta)$.
(b) The lemniscate $r^2 = 8\cos 2\theta$.

**10.** Find the points of intersection of each pair of curves. Be sure you get *all* intersections.
(a) $r^2 = 4\sin 2\theta$, $r^2 = 4\cos 2\theta$.
(b) $r = 2\sqrt{3}\cos\theta$, $r = 2\sin\theta$.
(c) $r = -4\cos\theta$, $r = -4\sqrt{3}\sin\theta$.
(d) $r = \sqrt{2}\sin\theta$, $r^2 = \cos 2\theta$.
(e) $r^2 = 4\sin\theta$, $r = 1 + \sin\theta$.

## 12-2 Parabolas, Ellipses, and Hyperbolas

*Parabolas*

There is some interest in using polar coordinates with parabolas. Take the origin at the focus of the parabola, and let the directrix be the line

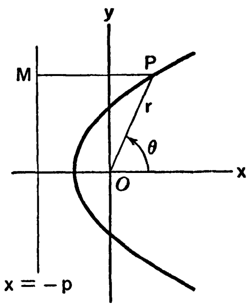



**Fig. 12-6**

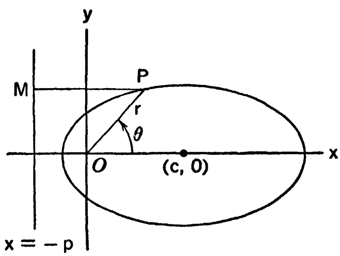



**Fig. 12-7**

$x = -p$, where $p > 0$ (see Fig. 12-6). Then $MP = p + r\cos\theta$, and so the definition of the parabola yields the equation $p + r\cos\theta = r$, or

$$r = \frac{p}{1 - \cos\theta}. \tag{1}$$

If the parabola is turned counterclockwise through an angle $\alpha$, keeping the focus fixed, $\cos(\theta - \alpha)$ takes the place of $\cos\theta$ in (1).

*Ellipses*

Suppose $p > 0$, and consider the locus of a point $P$ which moves so that its distance $OP$ from $O$ and its distance $MP$ from the line $x = -p$ are in constant ratio $e$, where $0 < e < 1$. This turns out to be an ellipse of eccentricity $e$ with $O$ as one focus, as we shall show. See Fig. 12-7. The defining relation is

$$\frac{OP}{MP} = \frac{r}{p + r\cos\theta} = e, \tag{2}$$

or

$$r = \frac{ep}{1 - e\cos\theta}. \tag{3}$$

Now (2) can be written in the form

$$\frac{\sqrt{x^2 + y^2}}{p + x} = e. \tag{4}$$

If we square and clear fractions we obtain

$$(1 - e^2)x^2 + y^2 - 2e^2px = e^2p^2. \tag{5}$$

Since $0 < 1 - e^2$, we recognize this as an equation of an ellipse. Moreover, (5) is equivalent to (4), for if $(x, y)$ satisfies (5) it must satisfy either (4) or the equation

$$\sqrt{x^2 + y^2} = -e(p + x).$$

But this equation would require $p + x < 0$, and would mean that with $P$ on the left of the line $x = -p$ we have $OP = ePM$, which is clearly impossible.

Starting from (5), we can find the center of the ellipse by completing the square in $x$. The result is that the equation can be brought to the form

$$\left(x - \frac{e^2p}{1 - e^2}\right)^2 + \frac{y^2}{1 - e^2} = \left(\frac{ep}{1 - e^2}\right)^2. \tag{6}$$

Now let us define $a$ and $b$ by the formulas

$$a = \frac{ep}{1 - e^2}, \qquad b = \frac{ep}{\sqrt{1 - e^2}}. \tag{7}$$

Observe that $a > b$. Then, in the usual notation for ellipses (see § 3-8), let

$$c = \sqrt{a^2 - b^2} = \frac{e^2p}{1 - e^2}. \tag{8}$$

We see that equation (6) now takes the form

$$\frac{(x - c)^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Hence our locus is an ellipse with center at $(c, 0)$. Since $c$ is the distance from a center to a focus, this means that the origin is a focus.

From (7) and (8) we find that $c/a = e$; this means that the constant ratio $e$ in (2) is the eccentricity of the ellipse. Thus we have a new geometric way of defining an ellipse. The line $x = -p$ is called the *directrix* of the ellipse corresponding to the focus $O$. By symmetry it is clear that there is another directrix associated with the other focus.

*Hyperbolas*

There is also a focus-directrix characterization of hyperbolas. We proceed just as in the case of the ellipse, except that now we assume $e > 1$.

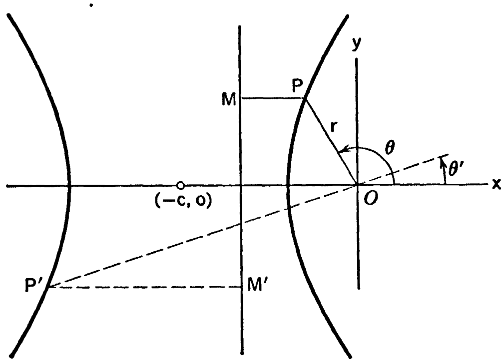



Fig. 12-8

We consider the locus of a point $P$ which moves in such a way that $OP/MP = e$, where $MP$ is the distance between $P$ and the line $x = -p$. See Fig. 12-8. The polar equation of this locus is

$$r = \frac{ep}{1 - e\cos\theta}. \tag{9}$$

On changing this equation to rectangular coordinates, squaring, simplifying, and introducing suitable notation, much as was done in the case of the ellipse, we obtain

$$\frac{(x+c)^2}{a^2} - \frac{y^2}{b^2} = 1,$$

where $$a = \frac{ep}{e^2 - 1}, \quad b = \frac{ep}{\sqrt{e^2 - 1}}, \quad c = \frac{e^2 p}{e^2 - 1}. \tag{10}$$

Hence the locus is a hyperbola of eccentricity $c/a = e$ with center at $(-c, 0)$ and one focus at $O$.

For values of $\theta$ such that $\cos\theta < 1/e$, equation (9) gives positive values of $r$. The corresponding points form the right-hand branch of the hyperbola. The left-hand branch is obtained from values of $\theta$ for which $\cos\theta > 1/e$; the corresponding values of $r$ are negative. This is illustrated by the point $P'$ in Fig. 12-8, with negative $r'$ corresponding to the positive $\theta'$.

The line $x = -p$ is called a directrix of the hyperbola. There is another directrix, with equation $x = -2c + p$.

Parabolas, ellipses (including circles), and hyperbolas are called, collectively, *conic sections*, because they are obtainable as intersections of a right circular cone and a plane. Which of the three types of curves one gets from such an intersection depends on the angle which the plane makes with the axis of the cone. In order to get both branches of a hyperbola one must take a plane which cuts both nappes of the cone.

Conic section curves are of great interest in connection with the study of the motions of planets and satellites. It was Kepler who announced that each planet moves in an elliptical orbit with the sun as a focus. This is a physical approximation of a general principle which was worked out later by Newton, using the inverse-square law of gravitation. If a mass particle moves, subject only to the force of gravitation between it and a fixed mass particle, the path of the moving particle is a conic section with the fixed particle as a focus (or as center if the path happens to be a circle). Later on in the book we shall show how to prove this assertion.

## EXERCISES

**1.** (a) For the ellipse in Fig. 12-7, express $p$ in terms of $a$ and $e$.
(b) Show that the distance from the center of the ellipse to one of the directrices is $a/e$.

**2.** For the hyperbola in Fig. 12-8, show that the distance from the center to a directrix is $a/e$.

**3.** For the ellipse of Fig. 12-7 suppose $r = 4$ when $\theta = \pi/3$ and $r = 3$ when $\theta = 3\pi/2$. Find the center of the ellipse and the point on the ellipse nearest 0.

**4.** Write an equation in polar coordinates for the ellipse with eccentricity 12/13, the origin at a focus, and the line $y = -25/12$ as the corresponding directrix. Where is the center? Where are the ends of the major axis?

**5.** Sketch and identify each of the following curves.

(a) $r = \dfrac{3}{2 - \cos\theta}$. (d) $r = \dfrac{16}{5 + 3\cos\theta}$.

(b) $r = \dfrac{2}{1 + \cos\theta}$. (e) $r = \dfrac{9}{4 - 5\sin\theta}$.

(c) $r = \dfrac{4}{1 + \sin\theta}$. (f) $r = \dfrac{25}{12 + 13\cos\theta}$.

**6.** For the parabola in Fig. 12-6 find a value of $\theta$ for which $OP$ is of the same length as the latus rectum.

**7.** For the hyperbola of Fig. 12-8 show that, when the ray $OP$ is parallel to an asymptote, its length is one-fourth that of the latus rectum.

**8.** A focal chord of a conic section is a line segment through the focus with ends on the curve. If $d_1$ and $d_2$ are the lengths into which such a chord is divided by the focus, show that the sum of the reciprocals of $d_1$ and $d_2$ is the same for all chords.

9. A comet, moving in a parabolic orbit and getting nearer the sun $S$, is 60 million miles from it at position $P_1$. When it is in the symmetrical position $P_2$, 60 million miles from the sun but going away from it, the angle $P_1SP_2$ is 120°. How near does the comet come to the sun (two possibilities)?

10. For the parabola (1) show that there is a value $\theta = \theta_1$ such that $-\pi/2 < \theta_1 < 0$ and the corresponding $r = r_1$ satisfies the condition $r_1 = 2r_2$, where $r_2$ corresponds to $\theta_2 = \theta_1 + \pi/2$. Show that $\theta_1 = \pi/2 - 2\tan^{-1} 2$. If $r_1 = 40$, what is the smallest value of $r$?

## 12-3 Arc Length and Tangents

### *Arc Length*

Consider a curve with equation $r = f(\theta)$ in polar coordinates. Let $s$ denote arc length measured along the curve in a specified direction from a specified point, so that $s$ is a function of $\theta$. We assume that $f$ has a continuous derivative. Now we can regard the curve as being defined parametrically, with $x = r\cos\theta$, $y = r\sin\theta$, and $r = f(\theta)$, so that $x$ and $y$ are functions of $\theta$. We know that $ds^2 = dx^2 + dy^2$. Now

$$dx = -r\sin\theta\, d\theta + \cos\theta\, dr, \qquad dy = r\cos\theta\, d\theta + \sin\theta\, dr. \tag{1}$$

We square these expressions, add, and simplify. The result is

$$ds^2 = dr^2 + r^2\, d\theta^2. \tag{2}$$

By using this formula we can compute $s$ by integration.

**Example 1:** Set up the integral for the total length of the limaçon $r = \sqrt{3} - \sin\theta$ (see Fig. 12-4).

From the equation of the curve we compute $dr = -\cos\theta\, d\theta$. Therefore

$$\begin{aligned} ds^2 &= \cos^2\theta\, d\theta^2 + (3 - 2\sqrt{3}\sin\theta + \sin^2\theta)\, d\theta^2 \\ &= (4 - 2\sqrt{3}\sin\theta)\, d\theta^2. \end{aligned}$$

The total length is

$$L = \int_0^{2\pi} (4 - 2\sqrt{3}\sin\theta)^{1/2}\, d\theta.$$

This is not an elementary integral. It can be transformed into the form of a standard elliptic integral.

The formula for $ds$ in polar coordinates can also be used in connection with areas of surfaces of revolution, just as in § 11-4, or in connection with mass distribution on the curve $r = f(\theta)$, using the ideas of § 11-9.

**Example 2:** Let the circle $r = 2b\cos\theta$ (see Fig. 12-2) be thought of as a material wire with variable density $\sigma = 2\cos\theta$ ounces per foot. Find the total mass and center of mass.

Clearly $\bar{y} = 0$, by symmetry. We have

$$M = \int \sigma \, ds, \qquad M\bar{x} = \int x\sigma \, ds.$$

An easy calculation shows that $ds^2 = 4b^2 \, d\theta^2$. We shall measure $s$ from $\theta = -\pi/2$, increasing as $\theta$ increases. Then $ds = 2b \, d\theta$ and $x = r\cos\theta = 2b\cos^2\theta$, and so

$$M = 4b\int_{-\pi/2}^{\pi/2} \cos\theta \, d\theta, \qquad M\bar{x} = 8b^2\int_{-\pi/2}^{\pi/2} \cos^3\theta \, d\theta.$$

We may integrate from 0 to $\pi/2$ if we double the results. Using formula 107 from the Table of Integrals, we have

$$M = 8b, \qquad M\bar{x} = \frac{32b^2}{3}, \qquad \bar{x} = \frac{4b}{3}.$$

*Tangents*

Sometimes it is convenient to know how to find the angle (which we denote by $\psi$) between the tangent to a curve at $P$ and the line $OP$ produced through $P$. Let the tangent be directed in the same sense as that in which $s$ increases. See Fig. 12-9. The angle $\psi$ is defined as the counterclockwise

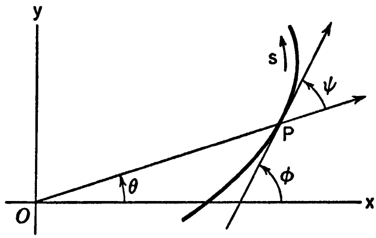



**Fig. 12-9**

angle from the directed ray $OP$ to the directed tangent at $P$. The angle $\phi$ is the inclination of the tangent. There is always a relation between $\theta$, $\phi$, $\psi$ of the form

$$\phi = \theta + \psi + n\pi, \tag{3}$$

where $n$ is some integer. In Fig. 12-9 it appears that $n = 0$, but there can be situations where $n \neq 0$. Once a definition choice of $\phi$, $\theta$, and $\psi$ has been made at one point of a curve, $n$ is determined, and the relation (3) will continue to hold as the point moves along the curves and the three angles vary continuously. Such a procedure may require the use of negative angles, or of angles greater than $2\pi$.

In order to obtain a formula for $\tan\psi$ in polar coordinates we proceed as follows. From (3) we see that

$$\tan\psi = \tan(\phi - \theta - n\pi) = \tan(\phi - \theta)$$
$$= \frac{\tan\phi - \tan\theta}{1 + \tan\phi\tan\theta}.$$

But $\tan \phi = dy/dx$ and $\tan \theta = y/x$, and so

$$\tan \psi = \frac{\dfrac{dy}{dx} - \dfrac{y}{x}}{1 + \left(\dfrac{dy}{dx}\right)\left(\dfrac{y}{x}\right)} = \frac{x\,dy - y\,dx}{x\,dx + y\,dy}.$$

Now $r^2 = x^2 + y^2$, and so $r\,dr = x\,dx + y\,dy$. From (1) we find that $x\,dy - y\,dx = r^2\,d\theta$. Hence, assuming that $r\,dr \neq 0$, we obtain

$$\tan \psi = \frac{r\,d\theta}{dr}. \tag{4}$$

At a point where $r \neq 0$ and $dr/d\theta = 0$, the tangent line is perpendicular to $OP$. This happens if $r$ attains a relative maximum or minimum.

The general relations between $\psi$, $ds$, $dr$, and $d\theta$ are shown in Fig. 12-10, where it is assumed that $r > 0$. The equations

$$\sin \psi = \frac{r\,d\theta}{ds}, \qquad \cos \psi = \frac{dr}{ds}, \tag{5}$$

which may be read from this figure, will be useful to us when we study motion of the point $P$ along the curve.

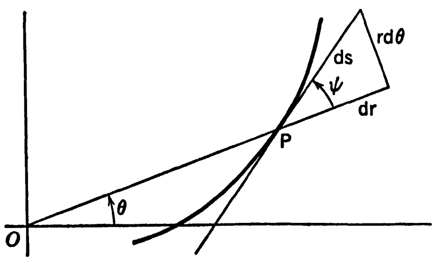



**Fig. 12-10**

If the curve passes through $O$ and has a tangent there, the limiting direction of $OP$, as $P$ approaches $O$ along the curve, is that of the tangent at $P$. This remark is of use when one is drawing the graph of a curve which goes through $O$. Some curves have a cusp at the origin. This can occur if $r$ is never negative, but approaches $O$, reaches it, and leaves it again as $\theta$ passes through a certain value $\theta_0$. In this case the ray $\theta = \theta_0$ will be tangent to the curve at the cusp. An example is furnished by the cardioid $r = a(1 + \cos \theta)$ at $\theta = \pi$.

As an example of the use of the formula for $\tan \psi$, we consider an interesting curve called the equiangular spiral.

**Example 3:** Consider the curve $r = ae^{k\theta}$, where $a > 0$ and $k \neq 0$. Suppose $k > 0$. Then it is clear that $r$ increases as $\theta$ increases. If we con-

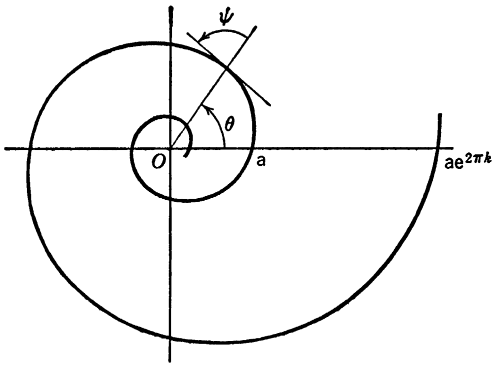



**Fig. 12-11**

sider all values of $\theta$, we see that $r \to +\infty$ as $\theta \to +\infty$ and $r \to 0$ as $\theta \to -\infty$. The remarkable feature of this curve is that $\psi$ is constant. We have

$$dr = kae^{k\theta}\, d\theta, \qquad \tan\psi = \frac{ae^{k\theta}\, d\theta}{kae^{k\theta}\, d\theta} = \frac{1}{k}.$$

The angle $\psi$ is $\tan^{-1}(1/k)$. See Fig. 12-11. If $k < 0$, the curve spirals inward instead of outward as $\theta$ increases.

## EXERCISES

**1.** Find $ds^2$ in terms of $\theta$ and $d\theta$ in each case.

(a) $r = b \sin\theta$.

(b) $r = a(1 - \cos\theta)$.

(c) $r(1 + \cos\theta) = a$.

(d) $r^2 = a^2 \sin 2\theta$.

(e) $r = 2\sin^2 \dfrac{\theta}{2}$.

(f) $r = 4\sin^3 \dfrac{\theta}{3}$.

**2.** Find the length of the spiral $r = 8e^{-\theta/2}$ from $\theta = 0$ to $\theta = 4\pi$.

**3.** Find the length of the spiral $r = e^{0.2\theta}$ from $\theta = -2\pi$ to $\theta = 2\pi$.

**4.** Find the total length of the cardioid $r = a(1 + \cos\theta)$.

**5.** Find the length of the indicated arc of each of the following curves. Do each problem in two ways, once integrating with respect to $\theta$, and once with respect to $r$.

(a) $r = 2\theta^2$ from $\theta = 0$ to $\theta = 3$.

(b) $r = \theta$ from $\theta = 1$ to $\theta = 2$.

(c) $r = 2/\theta$ from $\theta = \frac{1}{2}$ to $\theta = 4$.

(d) $r = 2\csc\theta$ from $\theta = 2\pi/3$ to $\theta = 3\pi/4$.

(e) $r = 4 \sin \theta$ from $\theta = \pi/4$ to $\theta = 2\pi/3$.
(f) $r = 4e^{\theta}$ from $\theta = 0$ to $\theta = \pi$.

**6.** Find the total length of the curve $r = a \sin^3 \dfrac{\theta}{2}$.

**7.** Find the area of the surface generated by revolving the cardioid $r = a(1 + \cos \theta)$ about the $x$-axis.

**8.** Find the area of the surface generated by revolving the lemniscate $r^2 = a^2 \cos 2\theta$ about the $y$-axis.

**9.** Locate the center of mass of the cardioid $r = a(1 + \cos \theta)$, thinking of it as a homogeneous wire.

**10.** Solve the preceding problem if the wire is not homogeneous, but has density $\sigma = r$.

**11.** (a) If the curves $r = f_1(\theta)$, $r = f_2(\theta)$ intersect (not at the origin) at a common value of $\theta$, show that they intersect orthogonally provided that $\tan \psi_1 \tan \psi_2 = -1$.
(b) Sketch the parabola $r(1 - \cos \theta) = a$ and the cardioid $r = a(1 - \cos \theta)$ and show that they intersect as right angles.

**12.** Show that the curves $r^2 = a^2 \sin 2\theta$, $r^2 = a^2 \cos 2\theta$ are orthogonal at their intersections where $r \neq 0$.

**13.** Find the angles at which each of the following pairs of curves intersect.
(a) $r = a \cos \theta$, $r = b \sin \theta$.
(b) $r = a \sin \theta$, $r = a(1 - \sin \theta)$.
(c) $r = \sqrt{2} \sin \theta$, $r^2 = \cos 2\theta$.

**14.** (a) For the upper half of the cardioid $r = a(1 + \cos \theta)$ show that $\psi = (\pi + \theta)/2$ and $\phi = (\pi + 3\theta)/2$ if we take $\psi$ and $\phi$ both equal to $\pi/2$ when $\theta = 0$.
(b) At what point is $\phi = \pi$? Check this by finding where $y$ is greatest on the cardioid.

## 12-4 Finding Area by Polar Coordinates

We now consider the problem of calculating the area which is swept out by the line segment $OP$ as $P$ moves from one point to another on the curve $r = f(\theta)$. Suppose the area is represented by $AOB$ in Fig. 12-12. This is the area swept out by $OP$ as $\theta$ increases from $\alpha$ to $\beta$. We shall express the area as an integral, and in order to do this we choose points $P_0, P_1, \cdots, P_n$ in order along the curve, with $P_0 = A$, $P_n = B$. The value of $\theta$ corresponding to $P_k$ is $\theta_k$ and $\theta_k - \theta_{k-1} = \Delta\theta_k$. We draw the rays from $O$ to the various $P_k$'s, thus dividing the area up into $n$ parts, and we proceed to obtain upper and lower approximating sums for the area. The basic formula for us here is the formula for the area of a circular sector. The area of a circular sector of radius $r$ and central angle $\Delta\theta$ is $\frac{1}{2}r^2 \, \Delta\theta$, for its area is the

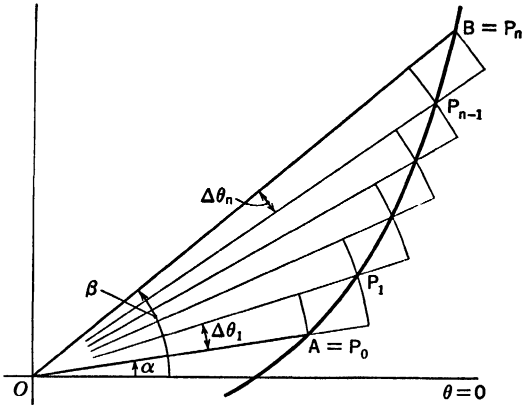



**Fig. 12-12**

fraction $\Delta\theta/2\pi$ of the total area of a circle of radius $r$. In Fig. 12-12 we now see that the required area is greater than

$$\tfrac{1}{2}(r_0^2\,\Delta\theta_1 + \cdots + r_{n-1}^2\,\Delta\theta_n)$$

and less than

$$\tfrac{1}{2}(r_1^2\,\Delta\theta_1 + \cdots + r_n^2\,\Delta\theta_n).$$

As $n$ increases and the maximum of the numbers $\Delta\theta_1, \cdots, \Delta\theta_n$ approaches zero, each of these sums approaches the integral

$$\frac{1}{2}\int_\alpha^\beta r^2\,d\theta \qquad [r = f(\theta)]$$

as its limit. Hence this integral yields the required area.

The foregoing discussion was based directly on the curve in Fig. 12-12, for which $r$ increases as $\theta$ increases. This is not always the situation, of course. In general, the area of the $k$th part of the total area is not less than $\frac{1}{2}m_k^2\,\Delta\theta_k$ and not more than $\frac{1}{2}M_k^2\,\Delta\theta_k$, where $m_k$ and $M_k$ are the smallest and largest values of $r$ for $\theta_{k-1} \le \theta \le \theta_k$. But the final result is still the same integral.

**Example:** Find the total area enclosed by the lemniscate $r^2 = a^2\cos 2\theta$ (see Fig. 12-5).

From symmetry it is evident that the total area is four times what we get as $\theta$ goes from 0 to $\pi/4$. Hence

$$A = 2\int_0^{\pi/4} a^2\cos 2\theta\,d\theta = a^2\sin 2\theta\Big|_0^{\pi/4} = a^2.$$

It is always essential to have a good notion of what the curve looks like in doing area problems, for the proper limits of integration will be determined by the figure. For example, in the foregoing problem it would not make sense to integrate from 0 to $\pi$, for there are no points on the curve corresponding to $\pi/4 < \theta < 3\pi/4$.

Now consider a point $P$ moving along the curve $r = f(\theta)$, starting from $\theta = \alpha$. Let $A$ be the area swept out by $OP$ from $\alpha$ to any given $\theta$, so that $A$ is a function of $\theta$. From our general area formula we see that

$$dA = \frac{1}{2} r^2 \, d\theta. \tag{1}$$

Hence, if we regard $A$ as a function of time $t$, we see that

$$\frac{dA}{dt} = \frac{1}{2} r^2 \frac{d\theta}{dt}. \tag{2}$$

The rate of change of $A$ figures in one of Kepler's laws of planetary motion, which asserts that a planet moves in such a way that the radius joining the planet to the sun sweeps out area at a constant rate. In other words,

$$r^2 \frac{d\theta}{dt} = \text{a constant} \tag{3}$$

for a given planet if we represent its orbit in polar coordinates with the sun at the origin.

## EXERCISES

**1.** Find the total area bounded by each curve.

(a) $r^2 = 16 \sin 2\theta$.
(b) $r = 2a \cos \theta$.
(c) $r = a(1 - \sin \theta)$.
(d) $r = 4a(1 + \cos \theta)$.
(e) $r = a \cos 2\theta$.
(f) $r = 8 \cos 3\theta$.
(g) $r = 4 \sin^2 \theta$.
(h) $r = a + b \cos \theta$, $a > b > 0$.

**2.** Find the area enclosed by one loop of each curve.

(a) $r^2 = 64 \sin \theta$.
(b) $r^2 = a^2 \cos 3\theta$.
(c) $r = a \cos n\theta$.
(d) $r = b \sin m\theta$.

**3.** Find the areas of the small loops of each of the following limaçons, and also the total area inside the outer part of the curve.

(a) $r = 4(1 + 2 \cos \theta)$.
(b) $r = 1 + 2 \sin \theta$.
(c) $r = 1 + \sqrt{2} \cos \theta$.
(d) $r = \sqrt{3} - 2 \sin \theta$.

**4.** Find the total area enclosed by the curve $r = 2 + \sin 2\theta$.

**5.** Find the area enclosed between the parabola $r(1 - \cos \theta) = p$ and the line on which $\cos \theta = 0$.

**6.** Find the area which is inside the circle $r = 2a \cos \theta$ but outside the circle $r = a$.

**7.** Find the area inside both the circle $r = a$ and the cardioid $r = a(1 + \sin \theta)$.

**8.** Show that the constant in (3) is $2S/T$, where $S$ is the area of the elliptical orbit and $T$ is the time required to go once around the orbit.

**9.** Suppose a point $P$ is tracing out the parabola $r(1 - \cos\theta) = p$ with decreasing $\theta$. Let $A$ be the area swept out by $OP$ in time $t$, starting when $\theta = \pi$, and suppose $dA/dt = k$, a positive constant. Show that $\theta$ and $t$ are related by the formula

$$\text{ctn}^3 \frac{\theta}{2} + 3\,\text{ctn}\,\frac{\theta}{2} = \frac{12kt}{p^2}.$$

## Review Questions and Problems for Chapters X, XI, and XII

### CONCEPTS AND DEFINITIONS

**1.** What difference, if any, is there between an indefinite integral and an antiderivative?

**2.** Upon what standard formula does the method of integration by parts depend?

**3.** What is meant by a rational function of two variables?

**4.** Write out a statement explaining how one defines the length of a curve. For what kind of a curve (i.e., for what particular form of analytical representation of a curve) is it possible to pass directly from the definition of the length of the curve to the formula for the length as a definite integral, without any use of Duhamel's principle?

**5.** What is the basic differential formula relating to length?

**6.** Define the total moment of a planar system of mass particles relative to an axis in the plane, and then define the center of mass. How is the definition extended to continuous distributions of mass?

**7.** What is meant by the centroid of a geometrical figure?

**8.** Does the concept of fluid pressure require a limit process for its precise elucidation? Explain.

**9.** Explain how to obtain *all* the polar coordinates of a point not at the origin. Illustrate for the point $x = 1$, $y = 1$.

**10.** Suppose a point $P$ on the curve $r = f(\theta)$ approaches the origin as $\theta \to \theta_0$. Explain why the ray $\theta = \theta_0$ is tangent to the curve at the origin.

### THEORY

**1.** If $f$ is continuous on $[a, b]$, how may a definite integral be used to provide a function $F$ whose derivative $F'(x)$ exists and is equal to $f(x)$ when $a \leq x \leq b$?

**2.** What important fact of algebra plays a central role in the systematic procedures for finding antiderivatives of rational functions?

**3.** Suppose that $R(s, t)$ is a rational function of $s$ and $t$.

(a) Suppose $f$ is a function of $x$ of the form $xR(x^2, \sqrt{a^2 - x^2})$. Explain a procedure for finding an indefinite integral of $f$ by methods of the text, without using trigonometric functions.

(b) Explain a procedure for reducing $\int f(x)\,dx$ to a form you know how to handle, if $f(x) = R(x, \sqrt{a^2 - x^2})$ or $f(x) = R(x, \sqrt{x^2 + a^2})$, where $R(s, t)$ is a rational function.

4. Use the law of the mean to pass from the definition of arc length to its expression as a definite integral for curves of the form $y = f(x)$. What assumption do you put on $f$?

5. Work out a justification of formula (1) or (2) in § 11-9 with more details than are given in the text, and show how the formula of Bliss comes into the work.

6. Derive the formula for $ds^2$ in polar coordinates.

7. Derive the formula $\tan \psi = r\, d\theta/dr$.

8. Apply the mean-value theorem to (7) in § 11-1 as applied to the arc length $\Delta s$ which comes with an increase of $\Delta t$ in the parameter $t$. Compare $\Delta s$ with the chord length $[(\Delta x)^2 + (\Delta y)^2]^{1/2}$, and prove that the ratio of $\Delta s$ to the chord length approaches 1 as $\Delta t \to 0$.

PROBLEMS

1. Find the indicated areas.
   (a) Between $y\sqrt{4 - x^2} = 8$ and the $x$-axis, from $x = -1$ to $x = \sqrt{2}$.
   (b) Between $y(9 + x^2) = 36$ and the $x$-axis, from $x = -\sqrt{3}$ to $x = 3$.
   (c) Between $y\sqrt{25 - 16x^2} = 4$ and $y = 1$.
   (d) Between the parabola $y^2 = 2(2 - x)$ and the $y$-axis.
   (e) Between the curve $(2x + 5)y = 10$ and the line $2x + y = 5$.

2. Work out the following indefinite integrals.

   (a) $\int \sin^3 2x \cos^4 2x\, dx$.  (d) $\int x \dfrac{\cos x}{\sin^2 x}\, dx$.

   (b) $\int \log(16 + x^2)\, dx$.  (e) $\int \dfrac{3 - x}{(9 + x^2)^{5/2}}\, dx$.

   (c) $\int \dfrac{\sin^{-1} x}{\sqrt{1 - x^2}}\, dx$.  (f) $\int \dfrac{\sin 2x}{\cos x + 4}\, dx$.

3. Find the volume generated when the area under one arch of $y = \sin x$ is revolved about the $x$-axis.

4. Find the area between one arch of the cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$ and the $x$-axis.

5. Find the volume generated when the area in Problem 4 is revolved around the $x$-axis.

6. Find the centroid of the arc cut off the parabola $y^2 = 4px$ by the latus rectum.

7. Find the centroid of the surface formed by revolving the arc of the preceding problem about the axis of the parabola.

**8.** Find the centroid of a homogeneous wire in the form of the catenary $y = 2 \cosh (x/2)$ from $x = 0$ to $x = 2$.

**9.** The part of the lemniscate $r^2 = 2a^2 \cos 2\theta$ on the right of the $y$-axis is revolved about the $x$-axis. Locate the centroid of the resulting surface.

**10.** For an ellipse represented in the standard way in polar coordinates, show that the mean value of $r$ with respect to $\theta$ $(0 \leq \theta \leq 2\pi)$ is the length of the semiminor axis.

**11.** A line segment of length $2b$ moves in the first quadrant with its two ends upon the $x$-axis and $y$-axis, respectively. Let $P$ be the foot of the perpendicular dropped from $O$ onto this segment. Show that the locus of $P$ is one loop of the curve $r = b \sin 2\theta$.

**12.** For the parabola $r(1 - \cos \theta) = p$ show that, if $\psi = \pi/2$ and $\phi = 3\pi/2$ when $\theta = \pi$, then $\psi = \pi - \dfrac{\theta}{2}$ and $\phi = \pi + \dfrac{\theta}{2}$ when $0 < \theta < 2\pi$. Use these facts to prove the optical property of the parabola.

**13.** (a) At what point in the first quadrant on $r^2 = 8 \cos 2\theta$ is $\psi = 2\pi/3$? $\psi = 5\pi/6$?

(b) Show that $\psi = 2\theta + \dfrac{\pi}{2}$ when $0 \leq \theta \leq \pi/4$.

(c) When is $\phi = 3\pi/4$? $\phi = \pi$?

**14.** (a) Supposing $a > b > 0$, discuss conditions on $a$ and $b$ which will insure that the limaçon $r = a + b \cos \theta$ is convex, i.e., that no chord of the curve goes outside the curve.

(b) For the general case of $r = f(\theta)$ where $r > 0$ always, show that the condition that the tangent always turns counterclockwise as $\theta$ increases is expressible in the form $r^2 + 2r'^2 \geq rr''$, where primes denote differentiation with respect to $\theta$.

# CHAPTER XIII

# MOTION IN A CURVE

## 13-1 Vectors as Number Pairs and as Geometric Objects

Many things in geometry and physics are such that we can attach numbers to them as a measure of "how much" of the thing there is. Lengths, areas, and volumes are measurable in terms of numbers. So are mass, work, and potential energy. If we consider a point moving along in a straight line, we can discuss its position in terms of a coordinate, which may be positive or negative, and we can discuss its velocity and acceleration in terms of the first and second derivatives of this coordinate with respect to time. *Direction* is important in these discussions of motion, but since the motion is confined to one line, we can handle all questions about direction by the use of both positive and negative numbers.

As soon as we turn to the motion of a point which need not stay on one line, but may move about in a plane, we can no longer handle questions of direction merely by the device of sign. The position of a point, and its velocity and acceleration, are things which cannot be represented or measured adequately by single numbers. We need *pairs* of numbers. We are already accustomed to the use of pairs of numbers to locate a point. From one point of view it is a very easy matter indeed to explain how we shall represent the velocity of a point by a pair of numbers. Suppose the point $(x, y)$ is moving in the $xy$-plane. Then $x$ and $y$ are functions of $t$. We shall say that the pair $(dx/dt, dy/dt)$ is the velocity of the point. Likewise we shall say that the pair $(d^2x/dt^2, d^2y/dt^2)$ is the acceleration of the point. But we shall not leave the matter in this rudimentary state. Velocity and acceleration are such important things that we need to gain a much better insight about them than is to be had merely by thinking of them as number-pairs obtained by differentiation from the pair $(x, y)$. We want a tangible geometric meaning for velocity and acceleration, to match our tangible visualization of the point as a *geometric object* rather than as a mere number-pair.

This is where vectors enter the scene. Vectors are mathematical entities which can be thought of in various ways. From one point of view a vector (in the plane) is a number-pair. From another point of view it is a geometric object. The mathematical usefulness of vectors is largely in their amenability to algebraic manipulation. But a very important feature of the use of vectors is that we are enabled to portray relationships geometrically, with all the advantages of insight we get from visualization of the things we discuss.

Under certain conditions, then, number-pairs are called vectors. A vector as a geometric object must be related in a definite manner to its other existence as a number-pair. As long as we are talking about one chosen rectangular coordinate system in a plane, every number-pair can be interpreted as a vector, and vice versa. The relation between a vector **A** as a geometric object and the number-pair $(a, b)$ which represents it is just this: **A** is the line segment from the origin to the point $(a, b)$. It is customary to affix an arrowhead to the segment at $(a, b)$ and to call this the *tip* of **A**. See Fig. 13-1. If $(a, b)$ happens to be the origin $(0, 0)$, our vector is called the *zero vector*. In this case the geometric object is just the point $O$.

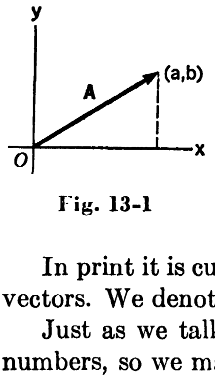



Fig. 13-1

In print it is customary to use boldface type for letters which represent vectors. We denote the vector corresponding to $(0, 0)$ by **0**.

Just as we talk about the *number system*, meaning the set of all real numbers, so we may talk about *the system of vectors in the xy-plane*. This system consists of **0** and all possible vectors such as **A** in Fig. 13-1, corresponding to all possible pairs $(a, b)$.

Now let us consider an example of a point moving in a simple way in the plane.

**Example:** Suppose that $x = t$, $y = \frac{1}{2}t^2$, where distances are in feet and time is in seconds. The vector for the moving point is $(t, \frac{1}{2}t^2)$. The velocity vector is $(1, t)$, and the acceleration vector is $(0, 1)$. Let us show these vectors on a diagram when $t = 2$. The point is moving on the curve $y = \frac{1}{2}x^2$, and at $t = 2$ the point is $(2, 2)$. We denote the vector to the point by **R**, the velocity by **V**, and the acceleration by **A**. Figure 13-2 shows the

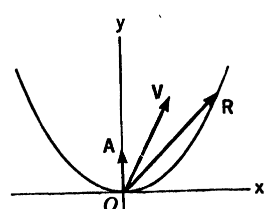



Fig. 13-2

situation. The vector **R** shows us the position of the moving point, and the vectors **V**, **A** show us the directions and magnitudes of the velocity and acceleration corresponding to this particular position of the point. It is these directions and magnitudes which are of physical interest.

The representation in Fig. 13-2 fails in one important respect to show us clearly something important about the velocity. *The velocity* **V** *is parallel to the straight line which is tangent to the curve at the tip of* **R**. *This is so in general, not merely in this particular case,* for we know by definition that **V** is represented by the pair $(dx/dt, dy/dt)$. Hence the slope of the vector **V** is

$$\frac{dy}{dt}\Big/\frac{dx}{dt} = \frac{dy}{dx},$$

and this, of course, is the slope of the curve.

The discovery of this fact suggests the following alteration in Fig. 13-2: Instead of drawing **V** and **A** as vectors issuing from 0, let us transport

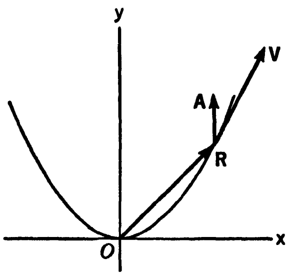



**Fig. 13-3**

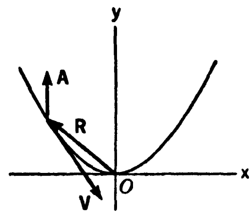



**Fig. 13-4**

them, without changing their magnitudes or directions, so that they issue from the tip of **R**. We call this *basing them* at the tip of **R**. See Fig. 13-3. Another position of the point (at $t = -\frac{3}{2}$) and the corresponding **V** and **A** are shown in Fig. 13-4.

This device of showing **V** and **A** based at the moving point instead of at $O$ is similar to the idea of showing heights of mountains on a topographical map by printing the height of a mountain adjacent to the spot which represents the top of the mountain, instead of keeping it as a number in its proper place on the number scale. It makes the information available visually in a self-explanatory way. In spite of this transporting of vectors from one place to another, the systematic development of rules for dealing with vectors in algebra and calculus is based on the concept of the system of all vectors as explained in connection with Fig. 13-1. *Vectors are directed line segments issuing from O.*

In the next section we discuss the algebra of vectors and differentiation of vector functions. After that we return to the study of velocity and acceleration.

## 13-2 Vector Algebra. Differentiation of Vector Functions

We shall discuss two kinds of algebraic operations with vectors. One of these operations is that of adding two vectors to get another vector. The other operation is that of multiplying a vector by a number to get another vector.

*Addition*

Suppose $\mathbf{A}_1$ is the vector represented by $(a_1, b_1)$ and let $\mathbf{A}_2$ be represented by $(a_2, b_2)$. Then we define $\mathbf{A}_1 + \mathbf{A}_2$ as the vector corresponding to $(a_1 + a_2, b_1 + b_2)$. If neither $\mathbf{A}_1$ nor $\mathbf{A}_2$ is $\mathbf{0}$ and if they are not collinear, the geometric interpretation of addition is that shown in Fig. 13-5, where $\mathbf{A}_1 + \mathbf{A}_2$ is the diagonal of the parallelogram formed on $\mathbf{A}_1$ and $\mathbf{A}_2$. Another way of putting it is this: To find the tip of $\mathbf{A}_1 + \mathbf{A}_2$, transport $\mathbf{A}_2$ without changing its magnitude or direction until it issues from the tip of $\mathbf{A}_1$. In this position the tip of $\mathbf{A}_2$ is where the tip of $\mathbf{A}_1 + \mathbf{A}_2$ is to be taken. This explanation also applies when $\mathbf{A}_1$ and $\mathbf{A}_2$ are collinear.

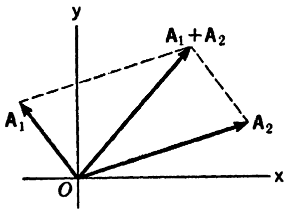


Fig. 13-5

In all cases $\mathbf{0} + \mathbf{A} = \mathbf{A} + \mathbf{0} = \mathbf{A}$.

We observe that addition is commutative and associative:

$$\mathbf{A}_1 + \mathbf{A}_2 = \mathbf{A}_2 + \mathbf{A}_1,$$

$$\mathbf{A}_1 + (\mathbf{A}_2 + \mathbf{A}_3) = (\mathbf{A}_1 + \mathbf{A}_2) + \mathbf{A}_3.$$

*Multiplication by a Number*

If $\mathbf{A}$ is a vector, represented by the pair $(a, b)$, and if $c$ is a number, we define $c\mathbf{A}$ as the vector which corresponds to the pair $(ca, cb)$. If either $c = 0$ or $\mathbf{A} = \mathbf{0}$, the product $c\mathbf{A}$ is $\mathbf{0}$. Otherwise the product is not $\mathbf{0}$, and the geometric interpretation of the product is this: If $c > 0$, $c\mathbf{A}$ is in the same direction as $\mathbf{A}$ and $c$ times as long. If $c < 0$, $c\mathbf{A}$ is in the direction opposite to that of $\mathbf{A}$ and it is $|c|$ times as long as $\mathbf{A}$.

The following algebraic rules are valid:

$$c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}, \qquad (c + d)\mathbf{A} = c\mathbf{A} + d\mathbf{A},$$

$$c(d\mathbf{A}) = (cd)\mathbf{A}, \qquad 1 \cdot \mathbf{A} = \mathbf{A}.$$

We agree that the factor $c$ can be written on either side of the vector: $c\mathbf{A} = \mathbf{A}c$.

The vector $(-1)\cdot\mathbf{A}$ is written $-\mathbf{A}$, and $\mathbf{B} + (-\mathbf{A})$ is written $\mathbf{B} - \mathbf{A}$. There is a simple geometric construction for $\mathbf{B} - \mathbf{A}$, resulting from the fact that $\mathbf{B} - \mathbf{A}$ is what must be added to $\mathbf{A}$ to give $\mathbf{B}$: Draw the line *from* the tip of **A** *to* the tip of **B**. The vector **B** − **A** has the same length and direction as this line. The student should construct a diagram to show this.

*Length of a Vector*

The length of a vector **A** is denoted by |**A**|. Observe that |**A**| > 0 except when **A** = **0**, and |**0**| = 0. Note also that |c**A**| = |c| |**A**|. If **A** is represented by $(a, b)$, then

$$|\mathbf{A}| = \sqrt{a^2 + b^2}. \tag{1}$$

*The Standard Unit Vectors*

The particular pairs (1, 0) and (0, 1) are called *the standard unit vectors*. We denote them by **i** and **j**:

$$\mathbf{i} = (1, 0), \qquad \mathbf{j} = (0, 1).$$

If **A** = $(a, b)$, we can express it as follows:

$$\mathbf{A} = a\mathbf{i} + b\mathbf{j}.$$

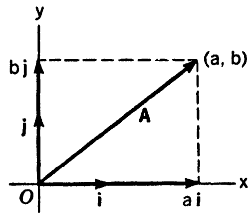



**Fig. 13-6**

We call $a$ the *x-component* of **A**. Sometimes we denote it by $A_x$. Likewise $b$ is the $y$-component, $A_y$. The vectors **i**, **j**, $a$**i**, $b$**j**, and **A** are shown in Fig. 13-6.

*Differentiation of Vector Functions*

We consider functions whose values are vectors; the domain of definition of the function is taken to be in the real number system. That is, we consider a law which sets up an assignment of a definite vector **R** corresponding to each number $t$ in the domain being considered. We shall think of $t$ as time, varying over some interval, or perhaps over the entire number scale. The location of the tip of **R** [call it $(x, y)$] is then determined by $t$, and as $t$ changes, $(x, y)$ varies. Since $x$ and $y$ are functions of $t$ (say $f$ and $g$), the study of our vector function is really equivalent to studying the parametric equations $x = f(t)$, $y = g(t)$. However, we are going to deal directly with the vector **R** and its changes as $t$ changes. We are going to use the Δ-notation as in § 3-1 and discuss the differentiation of **R** with respect to $t$.

If **R** corresponds to $t$ and $\mathbf{R}_1$ corresponds to $t + \Delta t$, where $\Delta t \neq 0$, let us write $\Delta\mathbf{R} = \mathbf{R}_1 - \mathbf{R}$, so that $\mathbf{R}_1 = \mathbf{R} + \Delta\mathbf{R}$. The derivative of **R** is defined as the limit

$$\frac{d\mathbf{R}}{dt} = \lim_{\Delta t \to 0} \frac{\Delta\mathbf{R}}{\Delta t}.$$

That is, we divide the vector Δ**R** by Δ$t$ and then find the limit of this new vector as $\Delta t \to 0$. The vector will approach a limit if and only if its tip approaches a limiting position. This will occur if and only if each of the components approaches a limit. In the present case $\mathbf{R} = (x, y)$, $\mathbf{R} + \Delta\mathbf{R} = (x + \Delta x, y + \Delta y)$,

$$\frac{\Delta\mathbf{R}}{\Delta t} = \left(\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}\right), \qquad \frac{d\mathbf{R}}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}\right).$$

However, we wish to visualize the process geometrically. If the tip of $\mathbf{R}$ follows a certain curve $C$ as $t$ varies, $\Delta\mathbf{R}/\Delta t$ will have the direction of the chord from the tip of $\mathbf{R}$ to the tip of $\mathbf{R}_1$, and hence $d\mathbf{R}/dt$ (if it is not $\mathbf{0}$) will have the direction of the tangent to the curve at the tip of $\mathbf{R}$, for the limiting direction of the chord is the direction of the tangent. See Fig. 13-7, in which we have transported the quotient vector and its limit away from their base at $O$ in order to see better the geometrical meaning of what is going on.

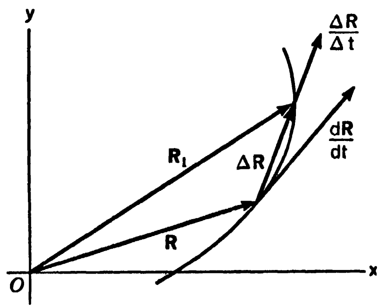



Fig. 13-7

Some rules about differentiation of vector functions can be developed. For us the most important rule is this: If a vector function is multiplied by a numerical function, and if both of them can be differentiated, then their product can be differentiated according to the rule

$$\frac{d}{dt}(u\mathbf{R}) = u\frac{d\mathbf{R}}{dt} + \frac{du}{dt}\mathbf{R}. \tag{2}$$

This is proved in just the same way that we proved the product rule for two numerical functions, in § 3-2. The rule for sums is just what one expects. Also, the derivative of a constant vector function is the vector $\mathbf{0}$. We shall see these rules put to use in succeeding sections of this chapter.

## EXERCISES

1. Find $2\mathbf{A} - 3\mathbf{B} + \mathbf{C}$ and its length if
   (a) $\mathbf{A} = 3\mathbf{i} - 2\mathbf{j}$, $\mathbf{B} = \mathbf{i} + 3\mathbf{j}$, $\mathbf{C} = 2\mathbf{i} + \mathbf{j}$;
   (b) $\mathbf{A} = -7\mathbf{i} + 4\mathbf{j}$, $\mathbf{B} = \mathbf{i} - \mathbf{j}$, $\mathbf{C} = -7\mathbf{i} + 7\mathbf{j}$.

   Make a diagram showing $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $2\mathbf{A} - 3\mathbf{B} + \mathbf{C}$ in each case.

2. Find a unit vector collinear with, but opposite in direction to, $-4\mathbf{i} + 3\mathbf{j}$.

3. If $\mathbf{A} = \sqrt{3}\,\mathbf{i} + \mathbf{j}$ and $\mathbf{B} = -\mathbf{i} + \sqrt{3}\,\mathbf{j}$, what angle does $\mathbf{A} + \mathbf{B}$ make with the positive $x$-axis? What angle does $\mathbf{A} - \mathbf{B}$ make? Draw a diagram.

4. Find a vector of length 26 and slope $-\frac{12}{5}$ (two answers).

**5.** Find a vector of length 1 which, if based at (4, 4) on $y^2 = 4x$, is normal to this curve there and points toward the positive $x$-axis.

**6.** (a) Show that $\frac{1}{2}(\mathbf{A} + \mathbf{B})$ extends from $O$ to the mid-point of the line segment joining the tips of **A** and **B**. (b) Where is the tip of $\frac{1}{3}\mathbf{A} + \frac{2}{3}\mathbf{B}$? What proportion of the way is it from the tip of **A** to the tip of **B**?

**7.** Show that as $t$ goes from 0 to 1, the tip of $(1 - t)\mathbf{A} + t\mathbf{B}$ goes along the line segment joining the tips of **A** and **B**, from the tip of **A** to the tip of **B**.

**8.** Find the vector from $O$ to the intersection of the medians of the triangle formed by $O$ and the tips of **A** and **B**.

**9.** Describe geometrically the locus of the tip of **R** if $\mathbf{R} = \mathbf{A} + t\mathbf{B}$ ($t$ variable), where neither **A** nor **B** is **0** and **B** is not collinear with **A**. Make a diagram.

**10.** What is the locus of the tip of **R** if $\mathbf{R} = a(1 - t)\mathbf{i} + bt\mathbf{j}$ ($a$ and $b$ nonzero constants)?

**11.** Show that the locus of the tip of $\mathbf{R} = t\mathbf{i} + (mt + b)\mathbf{j}$ ($m$ and $b$ fixed) is the line $y = mx + b$.

**12.** What is the locus of the tip of $\mathbf{R} = (t^2 + 2t + 2)\mathbf{i} + 2(t + 1)\mathbf{j}$?

**13.** Prove that **R** is perpendicular to $d\mathbf{R}/dt$ if

$$\mathbf{R} = \frac{2t}{1 + t^2}\mathbf{i} + \frac{1 - t^2}{1 + t^2}\mathbf{j}.$$

**14.** Express **R** as a function of $x$ if its tip moves on the curve $y = x^3$. Calculate $d\mathbf{R}/dx$ and $d^2\mathbf{R}/dx^2$ and draw them for $x = -1$; $x = 0$; $x = \frac{1}{2}$.

**15.** Find the locus of the tip of **R** if $\mathbf{R} = (3 \sin 4t)\mathbf{i} - (5 \cos 4t)\mathbf{j}$. Which way does it move on the curve as $t$ increases? Calculate $d\mathbf{R}/dt$ and $d^2\mathbf{R}/dt^2$ and show them on the diagram, with the curve, when $t$ is such that $x = \frac{9}{5}$, $y = 4$.

**16.** Find the locus of the tip of **R** if $\mathbf{R} = (2a \cos^2 t)\mathbf{i} + (a \sin 2t)\mathbf{j}$. Show $d\mathbf{R}/dt$ based at the tip of **R** for $t = 0$; $\pi/4$; $\pi/2$.

## 13-3 Vector Velocity

Consider a point moving on a curve in the $xy$-plane. The vector from the origin to the point is called the position vector of the point. We denote it by **R**. We can write $\mathbf{R} = x\mathbf{i} + y\mathbf{j}$, where the point is $(x, y)$. Throughout our discussion of motion in this chapter we assume that $x$ and $y$ have continuous first and second derivatives with respect to time. Since **i** and **j** are constant vectors, differentiation of **R** gives

$$\frac{d\mathbf{R}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j}. \tag{1}$$

This vector is called the velocity of the point. We denote it by **V**. Observe that the components of **V** are $dx/dt$ and $dy/dt$. Since we shall later consider components of other kinds, we call $dx/dt$ the $x$-component of $\mathbf{V}$, and denote it by $V_x$. Likewise $V_y = dy/dt$ is called the $y$-component of $\mathbf{V}$. It is a common practice to denote differentiation with respect to $t$ by placing a dot over a letter, so that $\dot{x} = dx/dt$. We shall sometimes use this notation.

The length of $\mathbf{V}$ is

$$|\mathbf{V}| = (\dot{x}^2 + \dot{y}^2)^{1/2}. \tag{2}$$

In view of the formula for $ds^2$ in § 11-1, we see that

$$|\mathbf{V}| = \left|\frac{ds}{dt}\right|. \tag{3}$$

In other words, the length of $\mathbf{V}$ is the speed at which the point is moving along the curve. We also know that if $\mathbf{V} \neq \mathbf{0}$, it has the direction of the tangent to the curve at the tip of $\mathbf{R}$ and is pointed in the same sense as that in which the particle is moving.

**Example 1:** Suppose $\mathbf{R} = (5 \cos 2t)\mathbf{i} + (3 \sin 2t)\mathbf{j}$. Discuss the path followed by the moving point. Find $\mathbf{V}$ as a function of $t$; find when $\mathbf{V}$ is longest and when it is shortest.

The curve has parametric equations $x = 5 \cos 2t$, $y = 3 \sin 2t$, so the path is the ellipse

$$\frac{x^2}{25} + \frac{y^2}{9} = 1.$$

The point $(x, y)$ goes around in the counterclockwise sense. The velocity is

$$\mathbf{V} = (-10 \sin 2t)\mathbf{i} + (6 \cos 2t)\mathbf{j}.$$

The speed (length of $\mathbf{V}$) is

$$|\mathbf{V}| = (100 \sin^2 2t + 36 \cos^2 2t)^{1/2} = (64 \sin^2 2t + 36)^{1/2}.$$

From this it is evident that the smallest speed is 6; it occurs when $\sin 2t = 0$, i.e., when $(x, y)$ is at either end of the major axis. The largest speed is 10; it occurs when $\sin 2t = \pm 1$, which is when $(x, y)$ is at either end of the minor axis.

In some problems we may not know $\mathbf{R}$ explicitly as a function of $t$, but we may be able to find $\mathbf{V}$ from the data given. The principle is this: if we know the equation satisfied by $x$ and $y$, and if we know *one* of the three quantities $\dot{x}$, $\dot{y}$, $\dot{s}$, we can find the other two and hence find $\mathbf{V}$.

**Example 2:** Suppose that a point moves along the parabola $y^2 = 2x$ at the rate of 3 feet per second ($x$ and $y$ in feet) and is going toward the vertex. Find $\mathbf{V}$ at the point $(2, 2)$. There are two possible methods.

*First Method.* Here we rely on similar triangles and on the fact that $\mathbf{V}$ can be pictured as a vector of length 3 based at $(2, 2)$ and pointing in the proper sense tangent to the parabola at $(2, 2)$. See Fig. 13-8. The slope at $(2, 2)$ can be calculated from the equation of the parabola. The slope is $\frac{1}{2}$. Hence the triangle with dotted sides shown in Fig. 13-8 is similar to the triangle shown in Fig. 13-9. Since it is clear that $V_x$ and $V_y$ are both negative, we have

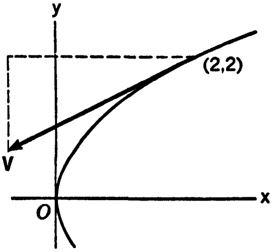



Fig. 13-8

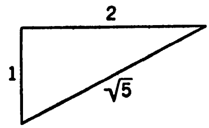



Fig. 13-9

$-V_x/3 = 2/\sqrt{5}$, $-V_y/3 = 1/\sqrt{5}$, or $V_x = -6/\sqrt{5}$, $V_y = -3/\sqrt{5}$. So, finally,

$$\mathbf{V} = -\frac{6}{\sqrt{5}}\mathbf{i} - \frac{3}{\sqrt{5}}\mathbf{j}.$$

*Second Method.* Here we start from the fact that $\dot{x}^2 + \dot{y}^2 = 9$, as a result of (2). From the equation of the parabola we find

$$2y\frac{dy}{dt} = 2\frac{dx}{dt}, \qquad \dot{x} = y\dot{y}.$$

Then

$$9 = (y^2 + 1)\dot{y}^2, \qquad \dot{y}^2 = \frac{9}{y^2 + 1}.$$

On substituting $y = 2$ and taking the square root, we find $\dot{y} = -3/\sqrt{5}$. We select the negative square root because we know from the given data that $\dot{y} < 0$. Finally, $\dot{x} = y\dot{y} = -6/\sqrt{5}$. We then write $\mathbf{V}$ just as in the first method.

## EXERCISES

In each exercise of this set, draw a figure showing the curve, and the vectors **R**, **V** for the particular instant in question, if one is specified.

1. For **R** as given or described, find general formulas for **V** and the speed of the moving point. Then answer the particular questions which are asked.

(a) $\mathbf{R} = (t^2 + 4)\mathbf{i} + (t - 2)\mathbf{j}$. What is the curve? Show the situation when the point crosses the $x$-axis. When is the speed least?
(b) $\mathbf{R} = (a \sin 2\pi t)\mathbf{i} + (2a \sin^2 \pi t)\mathbf{j}$. What is the curve? Describe the motion. Show the situation when $t = \frac{1}{3}$. What is the nature of the speed?
(c) $\mathbf{R} = \mathbf{i}t^2 + \mathbf{j}t^3$. What is the curve? How does it look near the point where the speed is zero? Show the situation when $t = 2$.
(d) $\mathbf{R} = 5(1 + \sin \pi t)\mathbf{i} + (4 \cos \pi t)\mathbf{j}$. What is the curve? Describe the motion. What is the periodicity? What is the maximum speed? The minimum speed? Show the situation at $t = \frac{4}{3}$.
(e) $\mathbf{R} = e^t\,\mathbf{i} + e^{-t}\,\mathbf{j}$. What is the path? Find the point where the speed is least.
(f) As in (e), if $\mathbf{R} = t^3\mathbf{i} + (24\sqrt{2} \log t)\mathbf{j}$ $(t > 0)$.

(g) As in (e), if $\mathbf{R} = t^2\mathbf{i} + (2t^2 - 25t)\mathbf{j}$. Show that as $t \to +\infty$, the limiting direction of $\mathbf{V}$ is perpendicular to the direction of $\mathbf{V}$ when the speed is least.

**2.** In each case a curve is given and there is a description of the motion at a certain instant. Find $\mathbf{V}$ at the instant in question, and diagram the situation.
(a) Point moving on $xy = -24$ in direction of increasing $y$, with speed 6 units per minute at $x = 4$.
(b) Point moving on $9x^2 + y^2 = 81$ with speed 10 and $x$ increasing at point $(-2\sqrt{2}, -\sqrt{3})$.
(c) Point moving on $2y = x^2$ with $V_y = -2$ at the point (4, 8).
(d) Point moving on $y^2 = 3x + 4$ with speed 3 and $y$ decreasing at (0, 2).

**3.** A point is on $4x^2 - y^2 = 64$, in the first quadrant and getting nearer the origin. Its speed is 9 units per second.
(a) Find the velocity when $y = 8$.
(b) How fast is the distance from the origin decreasing at this instant?

**4.** A point is at $(-8, 6)$ on $8(y + 2) = x^2$, with $x$ increasing and the distance from the origin decreasing at the rate of 2 units per minute. Find $\mathbf{V}$.

**5.** A point is moving on the parabola $y^2 = 2px$ with $y < 0$ and $\dot{x} < 0$. Let $v$ be the speed, $D$ the distance to the focus, and $r$ the distance to the origin. Show that: (a) $\dot{D} = \dot{x}$, (b) $v = \sqrt{2pD}\,|\dot{x}|/y$, (c) $\dot{r} = (x + p)\dot{x}/r$.

**6.** A boy is flying his kite, and paying out the string. (a) If the kite rises along the curve $9x^2 = 2{,}000y$ (the $y$-axis vertical, and the origin at the boy's feet), and if the horizontal velocity of the kite is 4 feet per second, find its vertical velocity and the speed in the path when the kite is 200 feet horizontally from the boy. (b) How fast is the boy paying out the string, on the assumption that it forms a straight line from the boy's feet to the kite?

**7.** A point $P$ moves on the parabola $y^2 = 4x$ with a speed of 2 units per minute. The tangent at $P$ intersects the line $x = -1$ in a point $Q$. Assuming $y > 0$, $V_x > 0$, find a general expression in terms of $x$ for the speed of $Q$. Evaluate when $P$ is at (1, 2).

**8.** A railroad track is curved in the shape of the parabola $y^2 = 1{,}000x$ ($x$ and $y$ in feet). A road is laid out along the $y$-axis. A night train, going 30 miles per hour, is approaching the vertex of the parabola. (a) How fast is the train approaching the road when the distance from the road is 1,000 feet? (b) How fast is the light from the train's headlight moving along the road?

**9.** If a rod 8 inches long, pivoted at one end, is lifted up to a horizontal position and released, the rate of change of the angle $\theta$ between the rod and the downward vertical is given by the formula $(d\theta/dt)^2 = 8 \cos \theta$. Find general expressions for the vertical and horizontal components of velocity of the free end of the rod, and evaluate them when $\theta = 60°$.

10. A bead placed on a smooth wire having the form of the curve $x^2y = 16$ (the $y$-axis being vertical) will slide down the wire with speed $v$ given by the formula $v^2 = 2g(y_0 - y)$, where $v = 0$ when $y = y_0$. Take $g = 32$, $y_0 = 16$, and find $V_x$ and $V_y$ when $y = 2$, assuming that $x$ is positive. What limit does $V_x$ approach as $y \to 0$?

## 13-4 Vector Acceleration

We continue the general discussion of the motion of a point along a curve in the $xy$-plane.

The derivative of $\mathbf{V}$ with respect to $t$ is called the acceleration vector $\mathbf{A}$:

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{d^2\mathbf{R}}{dt^2} = \ddot{x}\mathbf{i} + \ddot{y}\mathbf{j}. \tag{1}$$

The two dots indicate a second derivative with respect to $t$. The $x$ and $y$ components of $\mathbf{A}$ are denoted by $A_x$ and $A_y$, respectively.

The direction of the acceleration vector is not usually that of the tangent to the curve at the tip of $\mathbf{R}$. As we shall see later on, if $\mathbf{A}$ is based at the tip of $\mathbf{R}$, it usually extends onto the concave side of the curve there. In exceptional cases it may be tangent to the curve. This happens at a point of inflection.

A simple case of great interest is that in which a point travels in a circular path at constant speed.

**Example 1:** Suppose $(x, y)$ goes counterclockwise around the circle $x^2 + y^2 = a^2$ with constant speed $v$.

In this case $\dot{s} = a\dot{\theta} = v$, where $\theta$ is the polar angle. Hence $\dot{\theta} = v/a$, and $\theta = vt/a$ if $t = 0$ when $\theta = 0$. Hence the path can be represented parametrically in the form

$$\mathbf{R} = \left(a\cos\frac{vt}{a}\right)\mathbf{i} + \left(a\sin\frac{vt}{a}\right)\mathbf{j}.$$

The velocity is

$$\mathbf{V} = \left(-v\sin\frac{vt}{a}\right)\mathbf{i} + \left(v\cos\frac{vt}{a}\right)\mathbf{j},$$

and the acceleration is

$$\mathbf{A} = \left(-\frac{v^2}{a}\cos\frac{vt}{a}\right)\mathbf{i} + \left(-\frac{v^2}{a}\sin\frac{vt}{a}\right)\mathbf{j}.$$

Observe that $\mathbf{A} = -v^2\mathbf{R}/a^2$. This means that $\mathbf{A}$ is directed oppositely to $\mathbf{R}$. The length of $\mathbf{A}$ is

$$|\mathbf{A}| = \frac{v^2}{a^2}|\mathbf{R}| = \frac{v^2}{a}.$$

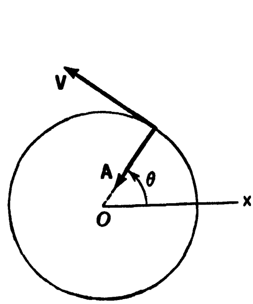


Fig. 13-10

The acceleration vector, if based at the tip of $\mathbf{R}$, points toward the center of the circle and has length equal to

$$\frac{(\text{speed})^2}{\text{radius}}.$$

The situation is shown in Fig. 13-10.

*Newton's Second Law*

In § 5-6 we discussed Newton's second law of motion as applied to mass particles moving on a straight line. The general form of the law is expressed in vector form by the equation

$$m\mathbf{A} = k\mathbf{F}, \tag{2}$$

where $m$ is the mass, $k$ is the same positive proportionality constant as in § 5-6, and the acceleration **A** and applied force **F** are now vectors.

When we look at the uniform circular motion of Example 1 from the point of view of Newton's law, we see that, in order to make a mass particle move in this way the force required must be of constant magnitude $mv^2/ka$ and must be directed toward the center of the circle.

Newton's law is in fact a vector differential equation (because $\mathbf{A} = d^2\mathbf{R}/dt^2$); we can attempt to solve it either by vector methods entirely, or else by studying the ordinary differential equations which result from examining various components of **A**. Most of such things are aside from our main pursuit just now.

*The Tangential Component of Acceleration*

Now consider once more the general case of a point moving along a curve. Let $s$ be arc length measured along the curve from some chosen point, and let the direction in which $s$ increases be considered the positive direction along the curve. The point itself may move in either direction.

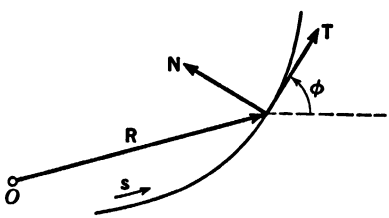


Fig. 13-11

Now think of **A** as being based at the tip of **R**. It is then possible to visualize **A** uniquely as the sum of two vectors, both based at the tip of **R**, one of them along the tangent to the curve, and the other at right angles to this tangent. In order to express this clearly, it is convenient to introduce two vectors of unit length, called **T** and **N**, which we define as follows: **T** is of unit length, has the direction of the tangent at the tip of **R**, and points in the direction of increasing $s$; **N** is of unit length, is at right angles to **T**, and is so directed that a 90° counterclockwise rotation brings **T** into the position of **N**. See Fig. 13-11. The vectors **T**, **N** are not constant, in general, since their directions may change as the point moves along the curve. If $\phi$ is the inclination of the tangent, we see that

$$\mathbf{T} = \mathbf{i}\cos\phi + \mathbf{j}\sin\phi, \qquad \mathbf{N} = -\mathbf{i}\sin\phi + \mathbf{j}\cos\phi. \tag{3}$$

Returning now to our discussion of **A**, let us express **A** as a multiple

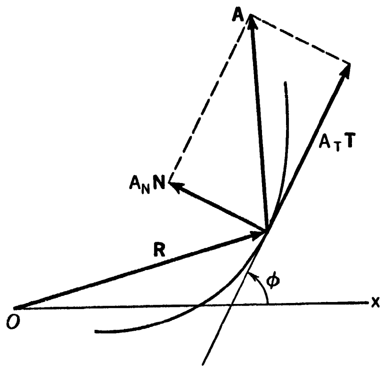



Fig. 13-12

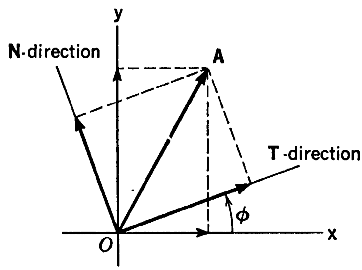



Fig. 13-13

of **T** plus a multiple of **N**. The necessary factors are denoted by $A_T$ and $A_N$, so that

$$\mathbf{A} = A_T\mathbf{T} + A_N\mathbf{N}. \tag{4}$$

The situation is shown in Fig. 13-12. Observe that, apart from sign, $A_T$ and $A_N$ are the lengths of the projections of **A** on the lines of **T** and **N**, respectively. We call $A_T$ and $A_N$ the tangential and normal components, respectively, of **A**.

We shall deal more fully with these components in the next section. Here we shall indicate one method of proving the formula

$$A_T = \frac{d^2s}{dt^2}. \tag{5}$$

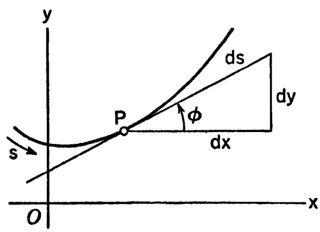



Fig. 13-14

If we go back to basing all our vectors at the origin, we can visualize the tip of **A** in two different ways (see Fig. 13-13). In the $xy$-coordinate system the tip of **A** is the point $\left(\frac{d^2x}{dt^2}, \frac{d^2y}{dt^2}\right)$. In the turned coordinate system, using the **T**-direction and **N**-direction, the tip of **A** is the point $(A_T, A_N)$. Now the angle of turning $\phi$ is just the angle of inclination of the tangent to the curve, and so the relations

$$\frac{dx}{ds} = \cos\phi, \qquad \frac{dy}{ds} = \sin\phi \tag{6}$$

hold, as a result of the relations between $dx$, $dy$, $ds$ (see Fig. 13-14). On the other hand, the rotation of axes formulas (2) in § 7-6, as applied here, show us that

$$\left.\begin{aligned} A_T &= \frac{d^2x}{dt^2}\cos\phi + \frac{d^2y}{dt^2}\sin\phi, \\ A_N &= -\frac{d^2x}{dt^2}\sin\phi + \frac{d^2y}{dt^2}\cos\phi. \end{aligned}\right\} \quad (7)$$

Hence, from (6), $$A_T = \frac{d^2x}{dt^2}\frac{dx}{ds} + \frac{d^2y}{dt^2}\frac{dy}{ds}. \quad (8)$$

Leaving this for a moment, consider the formula

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2.$$

Differentiating with respect to $t$, we obtain

$$2\frac{ds}{dt}\frac{d^2s}{dt^2} = 2\frac{dx}{dt}\frac{d^2x}{dt^2} + 2\frac{dy}{dt}\frac{d^2y}{dt^2}.$$

On multiplying this by $\frac{1}{2}\frac{dt}{ds}$, we get

$$\frac{d^2s}{dt^2} = \frac{dx}{ds}\frac{d^2x}{dt^2} + \frac{dy}{ds}\frac{d^2y}{dt^2}.$$

Comparing this with (8), we see that we have proved (5), as we set out to do.

Formula (5) is the actual means of computing $A_T$ in practice.

**Example 2:** Consider the point moving around the ellipse as described in Example 1, § 13-3. Find the vector acceleration and the tangential component of acceleration at $(4, \frac{9}{5})$.

From the earlier work we have

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = (-20\cos 2t)\mathbf{i} + (-12\sin 2t)\mathbf{j} = -4\mathbf{R}.$$

This means that the acceleration is always directed toward the center of the ellipse. Its magnitude is not constant, however, but is exactly four times the distance of the point from the center of the ellipse. At the point in question

$$\mathbf{A} = -16\mathbf{i} - \frac{36}{5}\mathbf{j}.$$

We also saw earlier that the speed of the point in its elliptical path is

$$\frac{ds}{dt} = (64\sin^2 2t + 36)^{1/2}.$$

We are assuming that $s$ is measured so that it increases as $t$ increases, so that

$ds/dt > 0$. We next find

$$A_T = \frac{d^2s}{dt^2} = \frac{1}{2}(64 \sin^2 2t + 36)^{-1/2} \cdot 128 \sin 2t \cdot 2 \cos 2t,$$

$$= \frac{64 \sin 2t \cos 2t}{(16 \sin^2 2t + 9)^{1/2}}.$$

At $(4, \frac{9}{5})$ we have $\sin 2t = \frac{3}{5}$, $\cos 2t = \frac{4}{5}$. Then

$$A_T = \frac{256}{5\sqrt{41}}.$$

## EXERCISES

**1.** Find $A_x$, $A_y$, and $A_T$ in each of the following cases, and diagram the situation for the particular conditions which are indicated.

(a) $\mathbf{R} = 4t\mathbf{i} + (64t - 16t^2)\mathbf{j}$ at $t = 1$; $t = 2$; $t = 4$.
(b) $\mathbf{R} = (t^2 + 4)\mathbf{i} + (t - 2)\mathbf{j}$ at $t = 2$.
(c) $\mathbf{R} = \mathbf{i}t^2 + \mathbf{j}t^3$ at $t = -2$.
(d) $\mathbf{R} = 3(1 + \cos \pi t)\mathbf{i} + 5(1 + \sin \pi t)\mathbf{j}$ at $t = \frac{5}{4}$.
(e) $\mathbf{R} = (2a \cos^2 t)\mathbf{i} + (a \sin 2t)\mathbf{j}$ at $t = 0$; $t = \pi/4$.
(f) $\mathbf{R} = \dfrac{2t}{1 + t^2}\mathbf{i} + \dfrac{1 - t^2}{1 + t^2}\mathbf{j}$ at $t = 0$; $t = 1$.

**2.** In each case a curve is given and there is a description of the motion at a certain instant. Find $\mathbf{A}$ at the instant in question, and diagram the situation. Also find the magnitude of $A_T$. Use the equation of the curve to find a general relation between $\dot{x}$ and $\dot{y}$. Then use this and the equation $\dot{s}^2 = \dot{x}^2 + \dot{y}^2$ as a basis for further work.
(a) On $y^2 = 8x$ with constant $V_x = 10$, at $(2, 4)$.
(b) On $4y = x^3$ with constant $V_y = 48$, at $(4, 16)$.
(c) On $y^2 = 3x + 4$ with constant speed 3 and $V_y < 0$, at $(0, 2)$.

**3.** A baseball is thrown so that it travels in the parabolic path $y = x - (x^2/200)$, with the constant horizontal velocity component $V_x = 40\sqrt{2}$ feet per second. (Compare with Example 2, § 5-7.) Make a diagram showing the velocity and acceleration at a typical instant. Express $x$ and $y$ as functions of the time in seconds, assuming $t = 0$ when $x = 0$. What is the initial tangential component of acceleration, assuming that $s$ increases as $t$ increases?

**4.** A circle of radius 2 feet rolls along on the upper side of the $x$-axis, making one revolution every 2 seconds. Study the motion of that point fixed on the circle which is at $(0, 0)$ when $t = 0$. Express its acceleration vector as a function of $t$, and show that $A_T = 2\pi^2 \cos(\pi t/2)$. Show that the acceleration vector is of constant length $2\pi^2$ feet and that it turns in the clockwise sense with constant angular velocity. This problem should be considered in connection with the information about the cycloid contained in § 5-8. If $\mathbf{R}$ is the position vector of the moving point and $\mathbf{R}_1$ is the position vector of the center of the rolling circle, show that the acceleration vector is $\mathbf{A} = \pi^2(\mathbf{R}_1 - \mathbf{R})$. What does this mean, geometrically?

5. A ladder $2a$ feet long is standing straight up against a wall. Then the foot of the ladder slides along the level ground away from the wall in such a way that the angle $\theta$ between the ladder and the wall increases at a constant rate. It takes 10 seconds for the ladder to reach a horizontal position. Find $\mathbf{V}$ and $\mathbf{A}$ for the mid-point of the ladder during this motion, as functions of $\theta$. What is the locus of the mid-point? Describe $\mathbf{A}$ geometrically.

## 13-5 Curvature

The notion of curvature of a curve comes from the observation that as a point moves along a curve with constant speed, the tangent line at the point is turning, and that the rate of its turning is somehow related to the sharpness or gradualness of the bending of the curve. In fact, a reasonable measure of "how much the curve is curving" is obtained directly from the angular velocity of the tangent. We now make this precise.

Let a positive direction along the curve be established. Let $s$ be arc length measured in this direction from a chosen point, and let $\phi$ be the counterclockwise angle from the positive $x$-axis to the positively directed tangent to the curve at the point corresponding to an arbitrary value of $s$. We can regard $\phi$ as a function of $s$. Then we define the *curvature* $K$ of the curve as

$$K = \frac{d\phi}{ds}. \tag{1}$$

If $\phi$ is measured in radians, the units of $K$ are radians per unit length. One could also speak of curvature as "so many degrees per hundred feet"; or still other units could be used.

The curvature may be either positive or negative, and it may be zero in certain cases. Since $K > 0$ means that $\phi$ is increasing as $s$ increases, it is clear that this makes the curve turn away to the left of the tangent as one advances along the curve, facing in the positive direction. A straight line is the only curve whose curvature is always zero.

We also define a number $R$ called the *radius of curvature*, by the formula

$$R = \frac{1}{|K|}. \tag{2}$$

This is on the presumption that $K \neq 0$. If $K = 0$, we do not define $R$. Note that $R$ cannot be negative or zero.

Calculation of $K$ is done by various formulas, depending upon how the curve is defined. The general case is that in which the curve is represented parametrically. If the parameter is $t$, and $x$ and $y$ are functions of $t$,

$$K = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{\pm(\dot{x}^2 + \dot{y}^2)^{3/2}}. \tag{3}$$

Observe that the denominator here is $\pm|ds/dt|^3$. The choice of sign is to be made in such a way that the denominator has the same sign as $ds/dt$.

The general derivation of (3) is considered in Exercise 7. We shall show how to derive the formula in the special case when $x$ is the parameter, so that the curve is represented by an equation $y = f(x)$. We start from the fact that $\tan\phi = y'$. Hence, taking differentials,

$$\sec^2\phi\, d\phi = dy' = y''\,dx, \qquad d\phi = \frac{y''\,dx}{1+\tan^2\phi}, \qquad d\phi = \frac{y''}{1+y'^2}\,dx.$$

Also, from $ds^2 = dx^2 + dy^2$ we have

$$ds^2 = (1+y'^2)\,dx^2, \qquad ds = \pm(1+y'^2)^{1/2}\,dx.$$

Here the sign of $ds/dx$ must conform to the facts in a particular situation. If $s$ is increasing as $x$ increases, $ds/dx > 0$, and we choose the plus sign; if $ds/dx < 0$, we choose the minus sign. Then

$$\frac{d\phi}{ds} = K = \frac{y''}{\pm(1+y'^2)^{3/2}}, \tag{4}$$

with the sign to be chosen as specified.

The question of sign does not arise in calculating the radius of curvature; we have, simply

$$R = \frac{(1+y'^2)^{3/2}}{|y''|}. \tag{5}$$

**Example 1:** Show that the radius of curvature of a parabola is least at the vertex.

We can choose the parabola in a position so that its equation is $x^2 = 2py$, $p > 0$. Then

$$y = \frac{1}{2p}x^2, \quad y' = \frac{x}{p}, \quad 1 + y'^2 = \frac{p^2+x^2}{p^2},$$

$$R = \frac{(p^2+x^2)^{3/2}/(p^2)^{3/2}}{1/p} = \frac{(p^2+x^2)^{3/2}}{p^2}.$$

From this it is clear that $R$ is smallest when $x = 0$, which is at the vertex. The minimal $R$ is $p$, which is the distance from the focus to the directrix.

### *The Normal Component of Acceleration*

We now wish to show the relevance of the concept of curvature to the discussion of acceleration by showing that

$$A_N = K\left(\frac{ds}{dt}\right)^2. \tag{6}$$

To prove this we use the vectors **T** and **N** which were introduced in § 13-4. We observe from equations (3) in § 13-4 that

$$\frac{d\mathbf{T}}{d\phi} = \mathbf{N}, \qquad \frac{d\mathbf{N}}{d\phi} = -\mathbf{T}. \tag{7}$$

From $\mathbf{R} = x\mathbf{i} + y\mathbf{j}$ we have

$$\frac{d\mathbf{R}}{ds} = \mathbf{i}\frac{dx}{ds} + \mathbf{j}\frac{dy}{ds}.$$

Using formulas (6) from § 13-4, we see that

$$\frac{d\mathbf{R}}{ds} = \mathbf{i}\cos\phi + \mathbf{j}\sin\phi = \mathbf{T}.$$

Now $$\frac{d\mathbf{R}}{dt} = \frac{d\mathbf{R}}{ds}\frac{ds}{dt}, \quad \text{and so} \quad \mathbf{V} = \frac{ds}{dt}\mathbf{T}.$$

Then $$\mathbf{A} = \frac{d\mathbf{V}}{dt} = \frac{ds}{dt}\frac{d\mathbf{T}}{dt} + \frac{d^2s}{dt^2}\mathbf{T}. \tag{8}$$

Here we are using rules of differentiation mentioned in § 13-2. We now write

$$\frac{d\mathbf{T}}{dt} = \frac{ds}{dt}\frac{d\phi}{ds}\frac{d\mathbf{T}}{d\phi} = \frac{ds}{dt}K\mathbf{N}.$$

Here we have used (7) and the definition of $K$. When this result is placed in (8), we obtain

$$\mathbf{A} = \frac{d^2s}{dt^2}\mathbf{T} + K\left(\frac{ds}{dt}\right)^2\mathbf{N}. \tag{9}$$

On recalling the definitions of $A_T$ and $A_N$, we see that (9) means

$$A_T = \frac{d^2s}{dt^2}, \qquad A_N = K\left(\frac{ds}{dt}\right)^2.$$

Thus we have derived (6) and given a new derivation of the formula for $A_T$ which was worked out in § 13-4.

**Example 2:** A train is going 88 feet per second (60 miles per hour) along a track which forms a parabola $x^2 = 1000y$ ($x$ and $y$ measured in feet). A 200-pound man, sitting on a smooth seat, slides over to the end of the seat next to the window which is on the side of the train toward the outside of the curve. Find the force with which the end of the seat presses on the man when the train is just passing through the origin of the parabola.

Let $s$ be measured along the parabola in the direction of increasing $x$, and suppose the train is going in this same direction. We consider the man as a mass particle moving on the parabola. Then

$$\frac{ds}{dt} = 88, \qquad \frac{d^2s}{dt^2} = 0.$$

Thus $A_T = 0$ in this case, and $\mathbf{A} = A_N\mathbf{N}$. We calculate the curvature from (4), using the plus sign in this case.

$$y' = \frac{x}{500}, \qquad y'' = \frac{1}{500}.$$

At $x = 0$ $\qquad K = \dfrac{1}{500}, \qquad A_N = \dfrac{(88)^2}{500} = 15.488.$

Now, from Newton's law, $m\mathbf{A} = k\mathbf{F}$. With mass and force both measured in pounds, $k = 32$ (see § 5-6). At $x = 0$ the tangent to the parabola is the $x$-axis. Hence at this point, the force $\mathbf{F}$ on the man has the direction of the positive $y$-axis and its magnitude is

$$\frac{mA_N}{k} = \frac{200(15.488)}{32} = 96.8 \text{ pounds.}$$

This is the force exerted by the end of the seat on the man, in the direction at right angles to the direction of motion of the train.

## EXERCISES

**1.** Find the curvature at the indicated point of the curve. Also, find where the radius of curvature is least.
(a) $y = x^3$ at $x = \frac{1}{2}$.
(b) $y = x^4$ at $x = 1$.
(c) $y = \sin x$ at $x = \pi/3$.
(d) $y = \log \cos x$ at $x = \pi/4$.
(e) $x = t^2 - 2t,\ y = 1 - 4t$ at $t = 1$.
(f) $x = 1 + \cos 2\pi t,\ y = 3 + 2 \sin 2\pi t$ at $t = \frac{1}{3}$.
(g) $x = 5 \sin t - 1,\ y = 2 \cos t + 3$ at $t = \pi/6$.

**2.** Find the radius of curvature of the ellipse $x = a \cos \theta,\ y = b \sin \theta$ at the end of each axis of symmetry.

**3.** Find a general expression for the radius of curvature in each case.
(a) $x = e^\theta \cos \theta,\ y = e^\theta \sin \theta$.
(b) $x = a(\cos \theta + \theta \sin \theta),\ y = a(\sin \theta - \theta \cos \theta)$.
(c) $x = a \log(\sec t + \tan t),\ y = a \sec t$.
(d) $x = a(\theta - \sin \theta),\ y = a(1 - \cos \theta)$.

**4.** Find the radius of a circle if the curvature is (a) 3 radians per foot; (b) 2 radians per hundred feet; (c) 34° per thousand yards.

**5.** A point moves along the curve $y = e^x$ with $V_x = 2$ units per second. How fast is the tangent line turning when $x = 0$?

**6.** Show that the radius of curvature of $y = a \cosh(x/a)$ at $(x, y)$ is $y^2/a$.

**7.** Derive formula (3) for $K$, using the idea by which (4) was derived.

**8.** Derive a formula for $K$ in terms of polar coordinates, as follows. Start with $\operatorname{ctn} \psi = (dr/r\, d\theta)$, from the discussion of $\psi$ in § 12-3. Show that

$$d\psi = \frac{(dr/d\theta)^2 - r\,(d^2r/d\theta^2)}{r^2 + (dr/d\theta)^2}\, d\theta.$$

Referring to Fig. 12-9, show that $K = (d\theta/ds) + (d\psi/ds)$, and hence derive the formula

$$K = \frac{r^2 + 2(dr/d\theta)^2 - r\,(d^2r/d\theta^2)}{\pm[r^2 + (dr/d\theta)^2]^{3/2}}.$$

**9.** As an alternative to the derivation of the formula for $K$ in Exercise 8, start from the formula (3) in the text, thinking of $x = r\cos\theta$, $y = r\sin\theta$, $r = f(\theta)$, and $\theta$ as the parameter. In this way $\psi$ is not involved.

**10.** These exercises are to be done using the formula for $K$ in Exercise 8. Find the radius of curvature in each case.
(a) Of $r = a\cos 3\theta$ at $\theta = 0$; $\theta = \pi/6$.
(b) Of $r = a(1 + \cos\theta)$ at $\theta = 0$; $\theta = \pi/2$.
(c) Of $r = e^\theta$ at $\theta = 0$.
(d) Of $r^2 = 2a^2\cos 2\theta$ at $\theta = \pi/6$.

**11.** Find the magnitude of $A_N$ for each motion at the point indicated.
(a) $x = 2t$, $y = 4 - t^2$, at $t = 2$.
(b) $x = 1/t$, $y = 4 - t$, at $t = 1$.
(c) $x = e^t$, $y = e^{-t}$, at $t = 0$.
(d) On $y = \log x$ with constant speed $v$ and $V_x > 0$, at $x = 1$.
(e) On $x^2 = 36y$ with constant $V_x = -12$, at $(18, 9)$.

**12.** Follow the directions of Exercise 11.
(a) $x = 3t^2$, $y = 2t^3$, at $t = 1$.
(b) $x = (1 - t)^2$, $y = (1 + t)^2$, at $t = 1$.
(c) On $xy + x = 1$ with $V_y = -2$, at $(1, 0)$.
(d) On $100y = x^3$ with $V_x = \frac{1}{2}$, at $x = 10$.
(e) On $x = a\cos^3\phi$, $y = a\sin^3\phi$ with constant speed $a/2$, at the point where $\phi = \pi/4$.

**13.** In a test of physiological reactions a 160-pound man is placed in a small vehicle which travels counterclockwise around the ellipse $3x^2 + y^2 = 432$ (units in feet and the positive directions of the $x$- and $y$-axes east and north, respectively) at the rate of 30 feet per second.
(a) Find the maximum and minimum forces experienced by the man in the direction normal to the ellipse.
(b) What is the normal force when he is traveling exactly northeast?

## 13-6 Velocity and Acceleration in Polar Coordinates

### *The Radial and Transverse Unit Vectors*

For some purposes it is convenient to talk about *radial* and *transverse* components of velocity and acceleration. In order to do this we introduce two new unit vectors, one having the direction of the polar radius from $O$ to the moving point, and the other one being at right angles to the first one, as shown in Fig. 13-15. The unit vector along $OP$ is called $\mathbf{u}_r$; that in the perpendicular direction 90° counterclockwise from $OP$ is called $\mathbf{u}_\theta$. It is clear from Fig. 13-15 that

$$\mathbf{u}_r = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta, \qquad \mathbf{u}_\theta = -\mathbf{i}\sin\theta + \mathbf{j}\cos\theta. \tag{1}$$

Hence
$$\frac{d\mathbf{u}_r}{d\theta} = \mathbf{u}_\theta, \qquad \frac{d\mathbf{u}_\theta}{d\theta} = -\mathbf{u}_r. \tag{2}$$

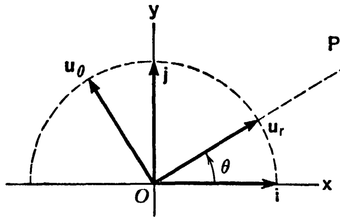


Fig. 13-15

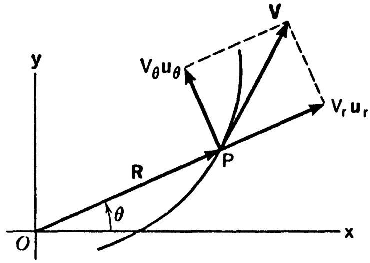


Fig. 13-16

*Velocity*

Let us now consider a point moving along a curve. We can express the velocity $\mathbf{V}$ as a multiple of $\mathbf{u}_r$ plus a multiple of $\mathbf{u}_\theta$:

$$\mathbf{V} = V_r\mathbf{u}_r + V_\theta\mathbf{u}_\theta. \tag{3}$$

The coefficients $V_r$, $V_\theta$ are called the radial and transverse components, respectively, of $\mathbf{V}$. See Fig. 13-16. We shall now obtain formulas for $V_r$ and $V_\theta$. For this purpose we express $\mathbf{R}$ in the form

$$\mathbf{R} = r\mathbf{u}_r. \tag{4}$$

For our present use of polar coordinates we assume that $r \geq 0$. From (4) we find

$$\frac{d\mathbf{R}}{dt} = r\frac{d\mathbf{u}_r}{dt} + \frac{dr}{dt}\mathbf{u}_r.$$

Now, in view of (2),

$$\frac{d\mathbf{u}_r}{dt} = \frac{d\mathbf{u}_r}{d\theta}\frac{d\theta}{dt} = \frac{d\theta}{dt}\mathbf{u}_\theta.$$

Hence

$$\mathbf{V} = \frac{d\mathbf{R}}{dt} = r\frac{d\theta}{dt}\mathbf{u}_\theta + \frac{dr}{dt}\mathbf{u}_r. \tag{5}$$

Comparing this with (3), we see that

$$V_r = \frac{dr}{dt}, \qquad V_\theta = r\frac{d\theta}{dt}. \tag{6}$$

We observe that $|\mathbf{V}| = (V_r^2 + V_\theta^2)^{1/2}$. This is just a different form of the equation

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2; \tag{7}$$

see (2) in § 12-3.

**Example 1:** A point $P$ moves counterclockwise around the limaçon $r = \sqrt{3} - \cos\theta$ in such a way that $OP$ turns at a constant rate, making 15 revolutions per minute. Find $V_r$, $V_\theta$, and the speed in the path as functions of $\theta$.

Here we know that $\dot{\theta} = 30\pi$ radians per minute. Also, $\dot{r} = (\sin\theta)\,\dot{\theta} =$

$30\pi \sin \theta$. Hence

$$V_r = 30\pi \sin \theta, \qquad V_\theta = 30\pi(\sqrt{3} - \cos \theta);$$

the speed in the path is

$$\frac{ds}{dt} = 30\pi(4 - 2\sqrt{3} \cos \theta)^{1/2}.$$

*Acceleration*

The radial and transverse components of acceleration are denoted by $A_r$ and $A_\theta$. To get formulas for them we start from (5) and differentiate.

$$\mathbf{A} = \frac{d\mathbf{V}}{dt} = r \frac{d\theta}{dt} \frac{d\mathbf{u}_\theta}{dt} + \frac{d}{dt}\left(r \frac{d\theta}{dt}\right) \mathbf{u}_\theta + \frac{dr}{dt} \frac{d\mathbf{u}_r}{dt} + \frac{d^2r}{dt^2} \mathbf{u}_r.$$

We have, from (2),

$$\frac{d\mathbf{u}_\theta}{dt} = \frac{d\mathbf{u}_\theta}{d\theta} \frac{d\theta}{dt} = -\frac{d\theta}{dt} \mathbf{u}_r.$$

A similar formula for the derivative of $\mathbf{u}_r$ has already been worked out. Thus

$$\mathbf{A} = -r\left(\frac{d\theta}{dt}\right)^2 \mathbf{u}_r + \frac{d}{dt}\left(r \frac{d\theta}{dt}\right) \mathbf{u}_\theta + \frac{dr}{dt} \frac{d\theta}{dt} \mathbf{u}_\theta + \frac{d^2r}{dt^2} \mathbf{u}_r.$$

The coefficient of $\mathbf{u}_\theta$ here is $A_\theta$; it can be written in the two forms

$$A_\theta = r \frac{d^2\theta}{dt^2} + 2 \frac{dr}{dt} \frac{d\theta}{dt} = \frac{1}{r} \frac{d}{dt}\left(r^2 \frac{d\theta}{dt}\right). \tag{8}$$

The coefficient of $\mathbf{u}_r$ is $A_r$. It is

$$A_r = \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2. \tag{9}$$

**Example 2:** Continuing the study of the motion in Example 1, we find

$$\ddot{r} = (30\pi \cos \theta)\, \dot{\theta} = (30\pi)^2 \cos \theta,$$

$$A_r = (30\pi)^2(2 \cos \theta - \sqrt{3}),$$

$$A_\theta = 2\dot{r}\dot{\theta} = (30\pi)^2\, 2 \sin \theta.$$

Observe that the acceleration is entirely radial when $\sin \theta = 0$. It is entirely transverse when $\cos \theta = \sqrt{3}/2$, which occurs at the points on the curve for which $x$ is largest. The magnitude of the acceleration is

$$|\mathbf{A}| = (A_r^2 + A_\theta^2)^{1/2} = (30\pi)^2(7 - 4\sqrt{3} \cos \theta)^{1/2}.$$

From this and earlier results it appears that $|\mathbf{V}|$ and $|\mathbf{A}|$ are largest at the same time, namely, when $\theta = \pi$. They are also smallest at the same time, namely, at $\theta = 0$. In both these cases $\mathbf{A}$ is perpendicular to the direction of motion.

*Central Forces and Kepler's Second Law*

Consider a mass particle at $P$ which moves in a plane, the motion being governed by a force acting on the particle in such a way that the force vector **F**, if based at the particle, points either toward the origin or away from the origin. For example, if there is a mass particle fixed at the origin, and if there are no other masses, the gravitational pull on the particle at $P$ will be directed toward the origin. A force always directed toward or away from $O$ is called a *central force.*

Now, in the case of a central force, Newton's law shows us that **A** also is directed either toward or away from the origin, and hence the transverse component of **A** is zero:

$$A_\theta = \frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) = 0.$$

As a result, $r^2(d\theta/dt)$ is constant. But from § 12-4 we see that this has the following consequence: The radius $OP$ sweeps out area at a constant rate. We see, therefore, that Kepler's second law about the motion of planets is a mathematical consequence of the law of gravitation, if we regard the motion of the planet as being essentially determined by the gravitational force between it and the sun.

In the chapter on differential equations, near the end of this book, we shall show how to prove that a mass particle which moves under the influence of gravitational attraction from just one other mass particle, which is fixed, moves in an orbit which is a conic section. The proof will begin with the formula we have derived for $A_r$.

## EXERCISES

**1.** In each case the motion of a point is described, and a particular location is specified. Draw a figure showing the curve, the vector **R** to the point in question, and the vectors **V**, **A** based at the point. The velocity and acceleration vectors are to be constructed after calculating their radial and transverse components.

(a) Counterclockwise on $r = 10 \sin \theta$ with speed 30 units per minute; at $\theta = \pi/4$.

(b) On $r = e^\theta$ with constant $\dot\theta = 2\pi$ radians per minute; at $\theta = 0$.

(c) On $r = a(1 + \cos \theta)$ with constant $\dot\theta = \pi$ radians per minute; at an arbitrary point, and then at $\theta = 0, \pi/2, \pi$.

(d) Counterclockwise on $r = a(1 + \cos \theta)$ with constant speed $a$ units per minute; at an arbitrary point, and then at $\theta = 0, \pi/2$. What happens to **A** as $\theta \to \pi$ with $\theta < \pi$?

(e) On $r = 2 + \sin 2\theta$ with constant $\dot\theta = 2\pi$ radians per minute; in general, and then at $\theta = \pi/4$, $\theta = \pi/2$, $\theta = 3\pi/4$.

**2.** Find $A_N$ and $A_T$ at the point indicated for each motion. Use the formula for $K$ in Exercise 8 of § 13-5.

(a) On $r = a(1 + \cos \theta)$ with $\dot\theta = 4\pi$ radians per minute; at an arbitrary point with $0 \le \theta \le \pi$.

(b) On $r = 3 + 2\cos\theta$ with $\dot{\theta} = \omega$ radians per minute; at an arbitrary point.
(c) On $r = a(1 - \sin\theta)$ with constant $V_r = a/3$; at $\theta = 0$ (with restriction $-\pi/2 < \theta < \pi/2$).
(d) On $r = 8\cos\theta$ with constant $V_\theta = 1$; at $\theta = \pi/8$ (with restriction $-\pi/2 < \theta < \pi/2$).

**3.** (a) For the parabola

$$r = \frac{p}{1 - \cos\theta} = \frac{p}{2}\csc^2\frac{\theta}{2},$$

with $s$ increasing as $\theta$ increases, show that

$$K = \frac{1}{p}\sin^3\frac{\theta}{2}\ (0 < \theta < 2\pi).$$

(b) If a point moves on this parabola, show that

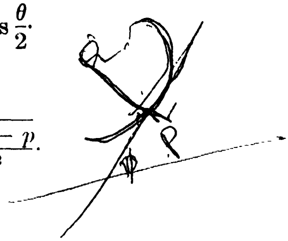
$$A_N = \frac{p\dot{\theta}^2}{4}\csc^3\frac{\theta}{2}.$$

(c) If the motion is such that $r^2\dot{\theta} = 2c$ (a constant), show that the speed is proportional to $r^{-1/2}$, and that

$$A_T = \frac{16c^2}{p^3}\sin^4\frac{\theta}{2}\cos\frac{\theta}{2}.$$

If $0 < \theta \leq \pi/2$ this can also be written

$$A_T = \frac{2\sqrt{2}\,c^2}{p}\frac{\sqrt{2r - p}}{r^{5/2}}.$$



## 13-7 The Center of Curvature

Let $P$ be a point tracing out a curve $C$ and suppose that, for the part of $C$ we consider, the curvature $K$ is never 0, so that the radius of curvature $R$ is well-defined. Since $K \neq 0$, the curve near $P$ lies entirely on one side of its tangent. Corresponding to $P$ we define a point $Q$ called the *center of curvature*, as follows. Construct the normal to $C$ at $P$ and proceed along it a distance $R$ from $P$ in the direction toward the concave side of the curve. The point $Q$ thus arrived at on the normal is called the center of curvature of $C$ corresponding to the point $P$. The circle with radius $R$ and center $Q$ is called the *osculating circle*. It is tangent to $C$ at $P$.

A very interesting concept is that of the *evolute* of a given curve $C$. We shall not take up this subject in great detail, but it deserves mention here. As $P$ moves along $C$, the locus of the corresponding center of curvature $Q$ is called the evolute of $C$. If $C$ happens to be a circle, the evolute is just one point, the center of $C$; but in general the evolute is another curve.

To find the center of curvature $Q$, we proceed as follows. Suppose $C$ is defined by an equation $y = f(x)$. If $P$ is $(x, y)$, denote the corresponding

$Q$ by $(X, Y)$. Then $Q$ lies on the line through $(x, y)$ with slope $-1/y'$, and therefore

$$Y - y = -\frac{1}{y'}(X - x).$$

Also, the distance $PQ$ is equal to $R$, and so

$$(X - x)^2 + (Y - y)^2 = R^2 = \frac{(1 + y'^2)^3}{y''^2}.$$

We solve simultaneously, eliminating $X$. The result is found to be

$$Y - y = \pm\frac{1 + y'^2}{y''}.$$

The ambiguity in sign is the expression of the fact that the normal cuts the circle in two points. We want the one which is on the concave side of $C$. By a diagram (which the student should make for himself) it can be seen that $Y - y$ and $y''$ should have the same sign. Hence

$$Y = y + \frac{1 + y'^2}{y''}. \tag{1}$$

Going back and solving for $X$ now, we find

$$X = x - \frac{y'(1 + y'^2)}{y''}. \tag{2}$$

Equations (1) and (2) give us the evolute in parametric form, with $x$ as the parameter. If $C$ is given in parametric form, these equations can still be used; all that is needed is to compute $y'$ and $y''$ in terms of the parameter.

**Example:** Find the evolute of the parabola $2py = x^2$.

From (1) and (2) we obtain, after simplification of the calculations for this case,

$$X = -\frac{x^3}{p^2}, \qquad Y = \frac{3x^2 + 2p^2}{2p}. \tag{3}$$

In this case we can eliminate the parameter and put the equation of the evolute in the form

$$27pX^2 = 8(Y - p)^3. \tag{4}$$

The appearance of the evolute in relation to the parabola is shown in Fig. 13-17.

There is another very interesting geometrical relationship between a curve and its evolute, and it can be visualized very clearly with Fig. 13-17 before us. The line $PQ$ is tangent to the evolute at $Q$. If we let $P$ move along the curve (with $x$ increasing in Fig. 13-17), *the length of $PQ$ increases at exactly the same rate as the arc length $Q_0Q$ along the evolute.* Hence we can think of the curve being traced out by $P$ as we unwind a string from the evolute, keeping the free part $QP$ tight as $Q$ moves and the string

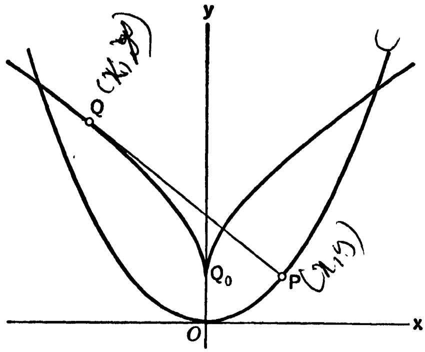



**Fig. 13-17**

winds or unwinds. Verification of this assertion, for the parabola in particular and for the case of an arbitrary curve, is left for the exercises.

## EXERCISES

**1.** Locate the center of curvature in each case.

(a) For $y = e^x$ at $x = 0$.

(b) For $xy = 4$ at $x = 2$.

(c) For $x^2y = a^2(x - y)$ at its maximum point.

**2.** Show that the evolute of the curve

$$x = a(\cos\theta + \theta\sin\theta), \qquad y = a(\sin\theta - \theta\cos\theta)$$

is the circle $x^2 + y^2 = a^2$. This curve is called the *involute* of the circle. It is generated by winding a string around the circle, leaving a free end at $x = a$, $y = 0$, and then unwinding the string counterclockwise, keeping the unwound portion straight. The length of the unwound section is $a\theta$. The free end generates the involute.

**3.** Obtain the evolute of $y^2 = x^3$ in terms of $x$ as a parameter. Plot several points on it and construct a diagram in the manner of Fig. 13-17.

**4.** Locate the centers of curvature corresponding to the ends of the axes of the ellipse $9x^2 + 25y^2 = 225$. Then sketch the evolute roughly, freehand.

**5.** For the cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$, show that parametric equations of the evolute are

$$X = a(\theta + \sin\theta), \qquad Y = -a(1 - \cos\theta).$$

The evolute is a cycloid also. It is generated by a circle of radius $a$ rolling on the line $y = -2a$. The relation between the original point $P$ and the corresponding center of curvature $Q$ can be visualized as follows: Draw the circle which generates the original cycloid, and an exactly equal circle tangent to it, but below the $x$-axis. Then $Q$ is on the lower circle, and the line $PQ$ passes through the point of tangency.

**6.** For the parabola and its evolute in Fig. 13-17, prove that $PQ$ is tangent to the evolute at $Q$, i.e., that $dY/dX = -1/y'$. Prove also that, if $S$ is the length of the arc $Q_0Q$ and $R = PQ$, then $dS/dx = dR/dx$. Then prove these same things in the general case, using the general formulas in the text.

*CHAPTER* XIV

# FURTHER STUDY OF LIMITS

## 14-1 The Purposes of this Chapter

In the very beginning of our study of calculus we encountered the concept of a limit when we defined a derivative:

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

In working out the details of finding derivatives of particular functions and developing general rules, we repeatedly found it necessary to consider limits. In a general sense, all these problems about limits could be posed in this way: Find the limit of a certain function $F(x)$ as $x$ approaches a certain value $x_0$. To begin with, the expression $F(x)$ might be a difference quotient, but nearly always it was transformed by algebraic or trigonometric manipulations into some simpler form, and the last part of the work consisted in using the rules about limits of sums, products, and quotients, as set forth in Theorems 1-C, 1-D, 1-E in § 1-8. These theorems were not proved at the early stage of Chapter I, even though they were needed for the logical continuation of our work. Now we have reached a point where it should be possible to go back and reconsider these theorems, with both a better perspective of their importance and a greater facility, born of experience, for understanding the discussion of such matters.

One purpose in this chapter, then, is to take up this unfinished business of the discussion of limits.

Another purpose is to discuss some fundamental aspects of the number system, and in connection with this discussion to consider the concepts of sequences and limits of sequences. This purpose is primarily related to preparation for the study of infinite series, in Chapter XV.

The last aim of the chapter is the exposition of a particular technique, known as *l'Hospital's rule,* for the finding of limits of functions in certain cases.

The three parts of the chapter are independent of each other, and need not be studied consecutively. However, the first part of § 14-2 should be read before reading § 14-3.

## 14-2 A Study of Inequalities. Proofs of the Limit Theorems

First of all, we must turn back to § 1-2 and read the remarks made there about inequalities and absolute values. For a thorough discussion of limits it will be necessary for us to go into more detail about inequalities and absolute values, and we shall now do this.

There are certain elementary algebraic rules for dealing with inequalities. These rules are easily understood and retained in mind by remembering that $a < b$ means "$b$ is to the right of $a$" on the real number scale, where the positive direction is to the right from 0. Thus $0 < x$ and "$x$ is positive" mean the same thing. The simplest rule about inequalities is this: If $a < b$ and $b < c$, then $a < c$.

If $a < b$ and $c$ is any number, then $a + c < b + c$. This rule permits us to transpose terms in inequalities just as in equalities. For example, from $a - 3 < x$ we obtain $a < x + 3$ by adding 3 on both sides.

If $a < b$ and $c$ is positive, then $ac < bc$. But if $a < b$ and $c$ is negative, then $ac > bc$ (the inequality is reversed).

Sometimes we wish to compare the sizes of fractions. We frequently use these evident facts: if $a$ and $b$ are positive, the fraction $a/b$ is made larger if we increase $a$, and it is also made larger if we *decrease* $b$.

The following facts about absolute values are important: for any numbers $a$, $b$,

$$|ab| = |a|\,|b|, \tag{1}$$

and

$$|a + b| \le |a| + |b|. \tag{2}$$

The correctness of these relations may be verified by considering the four cases: (1) $a$ and $b$ both positive, (2) both numbers negative, (3) one number negative and the other positive, (4) at least one of the numbers equal to zero. In cases (1), (2), and (4) it turns out that

$$|a + b| = |a| + |b|,$$

while in case (3) we have

$$|a + b| < |a| + |b|.$$

For three numbers we have

$$|a + b + c| \leq |a| + |b| + |c|,$$

and this extends to more than three in general by mathematical induction.

The ultimate basis of Theorem 1-C is the following assertion about sums: *In order to have $u + v$ differ from $u_0 + v_0$ by less than a specified positive number $k$, it is sufficient to have $u$ and $v$ differ from $u_0$ and $v_0$, respectively, by less than $k/2$.* This is a consequence of (2). For, if $|u - u_0| < k/2$ and $|v - v_0| < k/2$, then

$$|(u + v) - (u_0 + v_0)| = |(u - u_0) + (v - v_0)|$$
$$\leq |u - u_0| + |v - v_0| < \frac{k}{2} + \frac{k}{2} = k.$$

*Proof of Theorem 1-C.* We shall see that the foregoing assertion can be applied to yield a proof of Theorem 1-C. The student should now read Theorem 1-C and be ready to refer back to § 1-8 when necessary. Let us write $u = f(x)$, $v = g(x)$. With hypotheses as stated in Theorem 1-C we wish to show that we can make $f(x) + g(x)$ differ from $A + B$ by as little as we please, say by less than an assigned positive number $k$, simply by choosing some positive number $h$, (which we expect may depend on $k$) and insisting that $0 < |x - x_0| < h$. Now, the hypothesis that $f(x) \to A$ as $x \to x_0$ assures us that we can make $|f(x) - A| < k/2$ by insisting on a certain positive smallness for $|x - x_0|$; let us say that this is indicated by $0 < |x - x_0| < h_1$. Likewise $|g(x) - B| < k/2$ if $0 < |x - x_0| < h_2$. It is then clear that if we choose $h$ as the smaller of the numbers $h_1$, $h_2$ and insist on $0 < |x - x_0| < h$, we shall have both $|f(x) - A| < k/2$ and $|g(x) - B| < k/2$, and therefore $f(x) + g(x)$ will differ from $A + B$ by less than $k$, as required. This proves Theorem 1-C.

Next we consider an inequality problem connected with multiplication. If we know $u_0$ and $v_0$ and if we have an estimate of the size of $|u - u_0|$ and $|v - v_0|$, can we estimate the nearness of $uv$ to $u_0v_0$? A little trick of algebra will help us. We write

$$uv - u_0v_0 = u_0(v - v_0) + v_0(u - u_0) + (u - u_0)(v - v_0).$$

Then

$$|uv - u_0v_0| \leq |u_0|\,|v - v_0| + |v_0|\,|u - u_0| + |u - u_0|\,|v - v_0|.$$

Hence, if $|u - u_0| < p$ and $|v - v_0| < p$, we see that

$$|uv - u_0v_0| \leq (|u_0| + |v_0|)p + p^2.$$

This gives us a useful estimate. From it we see that, if $k$ is a given positive number, we can make $|uv - u_0v_0| < k$ by requiring $|u - u_0| < p$ and $|v - v_0| < p$, if $p$ is made small enough. For example, if $p$ is so small that $(|u_0| + |v_0|)p < \frac{1}{2}k$ and $p < (k/2)^{1/2}$, we shall have what is wanted.

*Proof of Theorem 1-D.* In Theorem 1-D let $u = f(x)$, $v = g(x)$, $u_0 = A$, $v_0 = B$, and take $p$ as indicated in the last sentence above. Then choose $h > 0$ so that, if $0 < |x - x_0| < h$, then $|f(x) - A| < p$ and $|g(x) - B| < p$. This can be done, since it is assumed that $f(x) \to A$ and $g(x) \to B$ as $x \to x_0$. The details of the preceding paragraph show that we then have $|f(x)g(x) - AB| < k$. This finishes the proof of Theorem 1-D.

Now consider quotients. We must avoid zero denominators, of course. Here our problem is to obtain an estimate of how near $u/v$ is to $u_0/v_0$ when we know how near $u$ and $v$ are to $u_0$ and $v_0$, respectively. We write

$$\frac{u}{v} - \frac{u_0}{v_0} = \frac{uv_0 - u_0v}{vv_0} = \frac{(u - u_0)v_0 + u_0(v_0 - v)}{vv_0},$$

$$\left|\frac{u}{v} - \frac{u_0}{v_0}\right| \leq \frac{|u - u_0|}{|v|} + \frac{|u_0|}{|v|\,|v_0|}\,|v - v_0|. \tag{3}$$

Now suppose that $|u - u_0| < p$ and $|v - v_0| < p$. In order to be safe about our denominator we shall assume that $0 < p < \frac{1}{2}|v_0|$. Then we can write $v_0 = v + (v_0 - v)$, and hence

$$|v_0| \leq |v| + |v_0 - v| < |v| + p < |v| + \tfrac{1}{2}|v_0|,$$

whence, on transposing $\frac{1}{2}|v_0|$, we find

$$\tfrac{1}{2}|v_0| < |v|.$$

Going back now to (3) and decreasing the denominators on the right by putting $|v_0|/2$ in place of $|v|$, we obtain

$$\left|\frac{u}{v} - \frac{u_0}{v_0}\right| < \frac{2p}{|v_0|} + \frac{2|u_0|}{|v_0|^2}\,p = 2\,\frac{|u_0| + |v_0|}{|v_0|^2}\,p.$$

It is clear from this discussion that, if $k$ is a given positive number, we can make

$$\left|\frac{u}{v} - \frac{u_0}{v_0}\right| < k$$

by choosing $p$ as a positive number smaller than both

$$\tfrac{1}{2}|v_0| \quad \text{and} \quad \frac{k|v_0|^2}{2(|u_0| + |v_0|)},$$

and then requiring that $|u - u_0|$ and $|v - v_0|$ be less than $p$.

*Proof of Theorem 1-E.* The proof of this theorem follows from the foregoing discussion in a manner quite like our proof of Theorem 1-D. With the same meanings of $u$, $v$, $u_0$, $v_0$ as in that proof, we choose $p$ as indicated above, and then choose $h$ so that, if $0 < |x - x_0| < h$, then $|u - u_0| < p$ and $|v - v_0| < p$. It will then follow that

$$\left|\frac{f(x)}{g(x)} - \frac{A}{B}\right| < k.$$

## 14-3 The Completeness Property of the Real Number System

A great deal of what we do in calculus, in so far as it depends on the nature of the real number system, is done by reliance on two general properties of the collection of all real numbers: (1) the properties embodied in the rules of addition, subtraction, multiplication, and division, including the special characteristics of the numbers 0, 1; (2) the properties embodied in the rules relating to inequalities. These rules were summarized in the beginning of § 14-2.

Numbers of the form $p/q$, where $p$ and $q$ are integers and $q \neq 0$, are called *rational* numbers. Numbers not of this sort are called *irrational.* It is not known exactly when or by whom the existence of irrational numbers was first discovered. But the irrationality of $\sqrt{2}$ (that is, the fact that there is no rational number $p/q$ such that $(p/q)^2 = 2$) was a known fact by some time in the latter part of the fifth century B.C., and the Greeks had developed a theory of incommensurables along geometrical lines. Now it is a fact that if we were to confine our attention exclusively to rational numbers, the system of rational numbers would exhibit all the properties described under (1) and (2) above. Hence there must be some further property of the system of real numbers beyond (1) and (2), a property which distinguishes the system of *all* real numbers from the system of rational numbers. There is indeed such a further property, and we shall proceed to explain what it is. For our explanation we must first introduce the concept of a *section* in the number system.

There are many ways of breaking the real number system up into two parts, a left-hand part $L$ and a right-hand part $R$, in such a way that each number in the part $L$ is less than each number in the part $R$.

**Example 1:** In $L$ put all negative numbers, and in $R$ put 0 and all positive numbers.

**Example 2:** In $L$ put all numbers less than or equal to $\sqrt{2}$, and in $R$ put all numbers greater than $\sqrt{2}$.

**Example 3:** In $L$ put all numbers $x$ such that there is some positive integer $n$ for which

$$x < \frac{3}{10} + \frac{3}{10^2} + \cdots + \frac{3}{10^n}, \tag{1}$$

and in $R$ put all numbers $x$ for which there is no such $n$. If $y$ is in $R$, then plainly

$$\frac{3}{10} + \frac{3}{10^2} + \cdots + \frac{3}{10^n} \leq y$$

for *every* $n$, and hence, if $x$ is in $L$, then $x < y$. We note that (1) is equivalent to $x < 0.33\cdots3$, where there are $n$ 3's to the right of the decimal point.

These are examples of what is called a section in the system of real numbers; $L$ is called the *lower part* of the section, and $R$ is called the *upper part*. The essential things required of a section are these: Every number must get put into either $L$ or $R$. If $x$ is put in $L$ and $y$ is put in $R$, then we must have $x < y$. And there must be some numbers (actually infinitely many) in each part of the section.

Now the real number system has this very important property: Whenever a section is made in the real number system, there is a unique number which is either the largest number in the lower part of the section or the smallest number in the upper part of the section. This property is called *completeness*. The unique number here referred to is called *the number determined by the section*.

In Example 1 the number determined by the section is 0, the smallest number in $R$. In Example 2 the number determined is $\sqrt{2}$, the largest number in $L$. It is not quite so easy to see what number is determined by the section in Example 3. The number is, in fact, the fraction $\frac{1}{3}$, and it belongs to $R$. For, $0.33\cdots3 < \frac{1}{3}$, no matter what finite number of 3's we have after the decimal point, and so $\frac{1}{3}$ and all numbers greater than $\frac{1}{3}$ are in $R$. And, if $x < \frac{1}{3}$, then when enough decimal places are taken, $0.33\cdots3$ differs from $\frac{1}{3}$ by less than the positive number $\frac{1}{3} - x$, and so $x < 0.33\cdots3$. Therefore all numbers less than $\frac{1}{3}$ are in $L$.

The completeness property of the real number system is at the root of several important theorems about continuous functions. Theorem 2-A is one of them. Another is Theorem 6-A. However, we shall not take up any further discussion of these theorems in this book. Our immediate purpose of bringing up the subject of completeness is to make it possible for us to give a satisfactory discussion of what are called *monotonic sequences*. There are two kinds of monotonic sequences, the nondecreasing ones and the nonincreasing ones.

Let $x_1, x_2, x_3, \cdots$ be an infinite succession of numbers such that $x_n \leq x_{n+1}$ for every positive integer $n$ (so that $x_1 \leq x_2 \leq x_3 \leq \cdots$). We say that the $x_n$'s form a nondecreasing sequence. In referring to the sequence we use the symbolism $\{x_n\}$ for the sequence as a whole. The $n$th member of the sequence, as an individual number, is $x_n$.

**Example 4:** Let $x_n = 2n + (-1)^n$. The first few terms are $x_1 = 1, x_2 = 5, x_3 = 5, x_4 = 9, x_5 = 9$. We observe that

$$x_{n+1} = 2(n + 1) + (-1)^{n+1} = 2n + 2 - (-1)^n,$$

and hence

$$x_{n+1} - x_n = 2 - 2(-1)^n.$$

This difference is 4 if $n$ is odd and 0 if $n$ is even. Thus certainly $x_{n+1} - x_n \geq 0$ in all cases, and the sequence is nondecreasing.

**Example 5:** Let $x_1 = 0.37$, $x_2 = 0.3737$, $x_3 = 0.373737$, and so on, with two more decimal places each time. In this case $x_1 < x_2 < x_3 < \cdots$.

There are two possibilities about the numbers $x_n$ in a nondecreasing sequence. Either there is a fixed number $A$ such that $x_n \leq A$ for every $n$, or there is no such $A$. In the first case we call $A$ *an upper bound* of the sequence and say that the sequence is *bounded.* In the second case we say that the sequence is *unbounded.* If there is an upper bound at all, there are many upper bounds. Thus, in Example 5, one upper bound is 4, and any number larger than 4 will also do. But 0.38 is likewise an upper bound, so is 0.374, and we could write down many more.

The sequence in Example 4 is unbounded. In fact, $x_n$ is $2n - 1$ if $n$ is odd and $2n + 1$ if $n$ is even, so that $x_n \geq 2n - 1$ in all cases, and the numbers $2n - 1$ clearly have no upper bound.

Another important type of sequence is the *nonincreasing* type: those sequences for which $x_n \geq x_{n+1}$ for every $n$ (so that $x_1 \geq x_2 \geq x_3 \geq \cdots$). Again there are two cases: either there is some fixed number $B$ such that $x_n \geq B$ for every $n$, or there is no such number. If there is such a $B$, we call it *a lower bound* and say that the sequence is bounded. If there is no lower bound we say that the sequence is unbounded.

**Example 6:**

$$x_1 = \frac{1}{2}, \quad x_2 = \frac{1\cdot 3}{2\cdot 4}, \quad x_3 = \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}, \quad \cdots, \quad x_n = \frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}.$$

In this case $x_1 > x_2 > x_3 > \cdots$. We get $x_{n+1}$ by multiplying $x_n$ by the positive factor $(2n + 1)/(2n + 2)$, which is less than 1. The sequence evidently has 0 as a lower bound.

We use the term *monotonic* to cover both the nondecreasing and the nonincreasing types of sequences.

If $\{x_n\}$ is a bounded nondecreasing sequence, there is a *smallest* upper bound of the sequence. The fact that there must be such a smallest upper bound can be shown clearly by constructing a suitable section in the number system and appealing to the completeness property. We make the section as follows: Into $R$ put all numbers $y$ such that $x_n \leq y$ for every $n$. That is, $R$ is composed of all the upper bounds of the sequence. Into $L$ we put all other numbers. It is easy (but essential) to check that this definition does indeed give us a section. Clearly all numbers less than $x_1$ go into $L$. If $x$ is in $L$ and $y$ is in $R$, then $x_n \leq y$ for *every* $n$ and $x < x_n$ for *some* $n$. Hence $x < y$. So we do have a section. Let $c$ be the number determined by the section. We shall prove that $c$ is the smallest number in $R$. Otherwise it would be the largest number in $L$. But this *cannot* be true, for if $c$ were in $L$, that would imply that $c < x_n$ for some $n$. Now let $b = \frac{1}{2}(c + x_n)$, so that $b$ is midway between $c$ and $x_n$. Then $b < x_n$, which implies that $b$ must be in $L$. Yet $c < b$, and we were supposing that $c$ was the largest number in $L$! This proves, then, that $c$, as the smallest number in $R$, is the smallest upper bound of the sequence. The usual term for $c$ is *the least upper bound* of the sequence.

Going back now to Example 5, let us ask: What is the least upper bound of the sequence $x_n$, where $x_n$ is the decimal $0.3737\cdots37$ (with $2n$ decimal places)? The answer is: the *nonterminating* repeating decimal $0.3737\cdots$, which, as we shall see later on (in Chapter XV) is the same number as the fraction $\frac{37}{99}$.

There is a parallel development of ideas for bounded nonincreasing sequences. In this case there is a *greatest lower bound* of the sequence. It can be obtained as the largest number in the lower part of a suitable section in the number system.

In the next section we shall consider sequences which may not be monotonic, and the general concept of the limit of a sequence. Then we shall see that every monotonic sequence which is bounded does have a limit. In the case of bounded nondecreasing sequences the limit is the least upper bound, and there is the corresponding situation for bounded nonincreasing sequences, where the limit is the greatest lower bound.

## EXERCISES

**1.** In § 2-6 we defined two sequences $\{A_n\}$, $\{S_n\}$ as follows:

$$A_n = \frac{8}{3}\left(2 - \frac{3}{n} + \frac{1}{n^2}\right), \qquad S_n = \frac{8}{3}\left(2 + \frac{3}{n} + \frac{1}{n^2}\right).$$

Show that these are monotonic sequences and determine the type of each. What least upper bounds or greatest lower bounds do you find here?

**2.** There are various convenient ways to test whether or not a sequence is monotonic. Sometimes this can be discerned merely by careful scrutiny of the general expressions for $x_n$ and $x_{n+1}$. In other cases it is convenient to express the difference $x_{n+1} - x_n$ as a single term and examine it. When $x_n$ is always positive, it is sometimes convenient to form the ratio $x_{n+1}/x_n$. If $x_{n+1}/x_n \geq 1$ for every $n$, the sequence is nondecreasing. Still another method which is sometimes applicable is this: Imagine $n$ to be a variable which is not restricted to integral values, and consider the derivative of $x_n$ with respect to $n$. If this derivative is negative when $n \geq 1$, the sequence is decreasing.

Apply any convenient method to decide as to whether each of the following sequences is monotonic, and if so, the type. Also investigate for boundedness or unboundedness, and say what you can about least upper bounds and greatest lower bounds.

(a) $x_n = \dfrac{n}{n+1}$.

(b) $x_n = \dfrac{n^2+1}{n^3}$.

(c) $x_n = n^2 - n$.

(d) $x_n = \dfrac{n^2}{n^3+n^2-1}$.

(e) $x_n = \left(\dfrac{n-2}{n+2}\right)^2.$ (g) $x_n = \dfrac{n^2+1}{n}.$

(f) $x_n = \dfrac{(n+1)^3 - n^3}{n^2}.$ (h) $x_n = \dfrac{(n+2)^2}{(n+1)!}.$

**3.** Some sequences are not monotonic from the very beginning, but become so after $n$ is sufficiently large. Discuss the following sequences with respect to such behavior.

(a) $x_n = n\left(\dfrac{3}{4}\right)^n.$ (d) $x_n = 1 - \dfrac{5^n}{n!}.$

(b) $x_n = \dfrac{n^2}{2^n}.$ (e) $x_n = \dfrac{n!}{100^n}.$

(c) $x_n = n(n+1)\left(\dfrac{2}{3}\right)^n.$ (f) $x_n = \dfrac{12n - 7n^2}{n^2 - 6}.$

**4.** Verify that the following definitions really make a section: Put $x$ into $L$ if $x \leq 0$; also put $x$ into $L$ if $0 < x$ and $x^2 \leq 2$. Put $y$ into $R$ if $0 < y$ and $2 < y^2$. What is the largest positive integer in $L$? What is the smallest positive integer in $R$? Without referring explicitly to any particular irrational number, demonstrate that if $x$ is positive and in $L$ and if $y$ is in $R$, then $x < y$. What is the number determined by this section, and in which part of the section is it?

**5.** If $\{x_n\}$ is a bounded nonincreasing sequence, define a section with $L$ and $R$ such that the greatest lower bound of the sequence is the largest number in $L$.

## 14-4 Convergent Sequences

By a *sequence* in general we mean an ordered infinite succession of numbers $x_1, x_2, x_3, \cdots$ determined according to some rule. This is equivalent to saying that $x_n$ is a function of $n$, the domain of definition of the function being the set of positive integers. We denote the sequence as a whole by $\{x_n\}$. A sequence need not be monotonic.

**Example 1:** (a) $x_n = (-1)^n/n$; (b) $x_n = \sin(n\pi/2)$. In the second of these sequences the terms go as follows: $1, 0, -1, 0, 1, 0, -1, 0, \cdots$, with the pattern repeating itself in blocks of four numbers.

A sequence $\{x_n\}$ is called bounded if there are two fixed numbers $A$, $B$ such that $A \leq x_n \leq B$ for every $n$. Otherwise the sequence is called unbounded.

Our main interest is in the limit concept for sequences. The definition is made in terms of inequalities. The sequence $\{x_n\}$ is said to have a certain number $c$ as a limit if for each positive number $\epsilon$ there is at least one corresponding positive integer $N$ such that

$$|x_n - c| < \epsilon \quad \text{if} \quad N \leq n. \tag{1}$$

When $c$ is related to $\{x_n\}$ in this way we write

$$\lim_{n\to\infty} x_n = c,$$

and say that $x_n$ converges to $c$. This is also expressed by saying that $x_n$ approaches $c$ as $n$ becomes infinite. For brevity this is often written in the manner

$$x_n \to c \quad \text{as} \quad n \to \infty.$$

Sometimes we curtail this simply to $x_n \to c$.

The meaning of (1) can be stated thus: $x_n$ will be as close to $c$ as we choose to require, provided merely that we insist upon $n$ being sufficiently large. The "sufficient largeness" of $n$ is expressed by the requirement $N \le n$. The size of $N$ will usually depend on the size of $\epsilon$.

If $\{x_n\}$ has a limit, the sequence is called *convergent.* A convergent sequence cannot have two different limits, because it is not possible for $x_n$ to be as near as we please to each of two different numbers for *all* sufficiently large $n$.

It is not ruled out that $x_n$ may be equal to its limit for some values of $n$, or even for infinitely many or all values of $n$.

**Example 2:** $x_n = \dfrac{1}{n}\sin\dfrac{n\pi}{2}\cdot$ Here the limit is 0. The first seven terms are $1, 0, -\frac{1}{3}, 0, \frac{1}{5}, 0, -\frac{1}{7}$, and this indicates the continuing pattern. We have $|x_n| < \epsilon$ if $\epsilon^{-1} < n$, so it suffices to take $N$ as the first integer larger than $\epsilon^{-1}$.

A convergent sequence is bounded, but not all bounded sequences are convergent. The sequence of Example 1(b) is an illustration of a bounded sequence which is not convergent.

Next we illustrate in two simple but important cases how the definition of a limit may be applied to show that a certain sequence does have a certain limit.

**Example 3:** $\lim\limits_{n\to\infty} \dfrac{4}{\sqrt{n}} = 0$. Here for a given $\epsilon > 0$, we wish to make $|4/\sqrt{n} - 0| < \epsilon$. This is equivalent, in turn, to

$$4/\sqrt{n} < \epsilon, \quad 4/\epsilon < \sqrt{n}, \quad 16/\epsilon^2 < n.$$

Hence, if we take for $N$ any integer larger than $16/\epsilon^2$ and require $N \le n$, we shall have the desired inequality.

Evidently an equally easy argument would show that

$$\lim_{n\to\infty} \frac{k}{n^p} = 0 \tag{2}$$

for any fixed $k$ and any $p > 0$.

**Example 4:** If $0 < r < 1$, then $\lim\limits_{n\to\infty} r^n = 0$. This is not quite so easy, though the result is certainly plausible. Let us introduce a new quantity $h$, defined by

$$h = \frac{1}{r} - 1, \quad \text{so that} \quad r = \frac{1}{1+h}.$$

Observe that $h > 0$, since $0 < r < 1$. Now, by the binomial expansion,

$$(1 + h)^n = 1 + nh + \text{positive terms},$$

and therefore $(1 + h)^n > 1 + nh$. Consequently

$$0 < r^n = \frac{1}{(1+h)^n} < \frac{1}{1+nh} < \frac{1}{nh}. \tag{3}$$

Now suppose that $\epsilon$ is any positive number. Let us choose $N$ so that

$$\frac{1}{Nh} < \epsilon, \quad \text{i.e.,} \quad \frac{1}{\epsilon h} < N.$$

Then $N \leq n$ will imply

$$\frac{1}{nh} \leq \frac{1}{Nh} < \epsilon,$$

and hence by (3) we shall have $|r^n - 0| < \epsilon$. This proves that $r^n \to 0$ as $n \to \infty$.

It is necessary for the student to acquire some familiarity with methods of finding the limits of sequences. Just as in the case of the limit concept for functions, as discussed in § 1-8 and § 14-2, so here, the consideration of sums, products and quotients is important. We have the following theorem:

THEOREM 14-A. *If $\{x_n\}$ and $\{y_n\}$ are convergent sequences with limits $a$ and $b$, respectively, then*

$$\lim_{n\to\infty} (x_n + y_n) = a + b, \qquad \lim_{n\to\infty} x_n y_n = ab,$$

*and with the additional hypothesis $b \neq 0$,*

$$\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{a}{b}.$$

This theorem can be proved at once on the basis of the discussion of inequalities in § 14-2. We omit details.

Next we consider an application of Theorem 14-A.

**Example 5:** Find $\lim_{n\to\infty} x_n$ if $x_n = \dfrac{3n^2 - 4n + 7}{2n^2 - n + 5}$. First we divide the numerator and denominator by the highest power of $n$ which occurs in the denominator:

$$x_n = \frac{3 - \dfrac{4}{n} + \dfrac{7}{n^2}}{2 - \dfrac{1}{n} + \dfrac{5}{n^2}}.$$

Then we apply (2) and Theorem 14-A to conclude that

$$\lim_{n\to\infty} x_n = \frac{3 - 0 + 0}{2 - 0 + 0} = \frac{3}{2}.$$

Sometimes we need to be able to recognize that a sequence is convergent, even though we cannot say precisely what number is the limit. In such a case we cannot use the definition of a limit directly. There are some useful methods for handling such situations. If the sequence is monotonic, we can rely on the following theorem.

THEOREM 14-B. *If a sequence is bounded and monotonic, it is convergent. If the sequence is nondecreasing, its limit is the least upper bound of the sequence. If the sequence is nonincreasing, its limit is the greatest lower bound.*

*Proof.* Suppose the sequence $\{x_n\}$ is nondecreasing, with $c$ as its least upper bound (see the discussion in § 14-3). If $\epsilon > 0$, we see that $c - \epsilon < c$. Hence $c - \epsilon$ is not an upper bound for the sequence ($c$ being the smallest upper bound) and there is some index, say $N$, such that $c - \epsilon < x_N$. Then, since the sequence is nondecreasing, $c - \epsilon < x_n$ if $N \leq n$. Since $x_n \leq c$ for every $n$, we then have $|x_n - c| < \epsilon$ if $N \leq n$, and so $x_n \to c$. The proof for nonincreasing sequences is similar.

**Example 6:** Let $x_n = 1 - (2^n/n!)$. The first few terms are $-1,\ -1,\ -\frac{1}{3},\ \frac{1}{3},\ \frac{11}{15}$. The sequence is nondecreasing, as a result of the fact that $2^n/n!$ decreases when $n$ increases, from $n = 2$ onward. In fact,

$$\frac{2^{n+1}}{(n+1)!} = \frac{2}{n+1}\frac{2^n}{n!} < \frac{2^n}{n!} \quad \text{if} \quad 2 \leq n.$$

It is also clear that $x_n < 1$ for every $n$. Hence $\{x_n\}$ is a bounded nondecreasing sequence. It must therefore be convergent, by Theorem 14-B. It is not part of our intent in this example to find the limit. But it is not hard to show that $2^n/n! < 4/n$ if $n > 2$, and from this it follows that $x_n \to 0$.

*Cauchy's Principle of Convergence*

There is a general theorem about convergent sequences which is useful because it applies to all convergent sequences, not merely monotonic ones. This theorem which we now state, is known as *Cauchy's principle of convergence*. It is named after the French mathematician Augustin-Louis Cauchy (1789–1857).

THEOREM 14-C. *A necessary and sufficient condition for the sequence $\{x_n\}$ to have a limit is that the absolute difference $|x_n - x_m|$ approach 0 as $m$ and $n \to \infty$.*

The meaning of the condition is that for each positive $\epsilon$ there is some positive integer $N$ such that $|x_n - x_m| < \epsilon$ if $N \leq m$ and $N \leq n$. We shall show how to prove that a sequence which satisfies this condition is indeed convergent. The proof of the converse, that a convergent sequence satisfies the condition, is left as an exercise.

We begin by constructing a section in the number system. A number $x$ is put into $R$ if there are infinitely many positive integers $n$ such that $x_n < x$. Otherwise we put $x$ into $L$. This means that $x$ is put into $L$ if there are *not* infinitely many $n$'s such that $x_n < x$. It is at once plain that if $x$ is in $L$ and $y$ is in $R$, then $x < y$, for the conditions which define $L$ and $R$ make $y \leq x$ impossible. We must show that there really are numbers which get put into $L$, and likewise for $R$. Using the hypothesis about the sequence, we choose $\epsilon = 1$ and let $N_1$ be an integer such that $|x_n - x_m| < 1$ if $N_1 \leq m$ and $N_1 \leq n$. Then, in particular, $|x_n - x_{N_1}| < 1$, or what is the same thing,

$$x_{N_1} - 1 < x_n < x_{N_1} + 1,$$

if $N_1 \leq n$. From this it follows at once that $x_{N_1} - 1$ must be in $L$ and $x_{N_1} + 1$ must be in $R$.

We now know that we have made a section in the number system. Let $c$ be the number determined by the section. We shall show that $x_n$ converges to $c$. First we observe the following fact: It is impossible for there to be two numbers $a$, $b$ such that $a < b$ and to have $x_n < a$ for an infinite number of values of $n$ and also to have $b < x_n$ for an infinite number of other values of $n$. This is because of our hypothesis about the sequence. For, since $b - a > 0$, there must be *some* integer $N$ such that $|x_n - x_m| < b - a$ for all indices $m$, $n$ except perhaps $1, 2, \cdots, N - 1$, and this makes it impossible to have an infinite number of terms of the sequence less than $a$ and also an infinite number greater than $b$. With this clearly in mind, we now suppose that $\epsilon$ is any given positive number. Our task is to show that there is some $N$ such that $c - \epsilon \leq x_n \leq c + \epsilon$ (which is equivalent to $|x_n - c| \leq \epsilon$) if $N \leq n$. Now $c - \epsilon$ is in $L$ and $c + \epsilon$ is in $R$, from the fact that $c$ is either the smallest number in $R$ or the largest number in $L$. Also, $c + \epsilon/2$ is in $R$. Hence we know the following: There are at most a finite number of $n$'s such that $x_n < c - \epsilon$, and there are an infinite number of $n$'s such that $x_n < c + \epsilon/2$. By the argument made earlier, then, there *cannot* be an infinite number of $n$'s such that $c + \epsilon < x_n$. Hence, except for some finite number of $n$'s, we have $c - \epsilon \leq x_n \leq c + \epsilon$. If we list the exceptional $n$'s and let $N - 1$ be the largest one (or else take $N = 1$ if there are no exceptions), we see that $|x_n - c| \leq \epsilon$ if $N \leq n$. This finishes the proof that $x_n$ converges to $c$.

The only use we make of Cauchy's principle of convergence in this book is in connection with the discussion of absolute convergence of infinite series, in § 15-5.

## EXERCISES

**1.** Find the limit of the sequence $\{x_n\}$ for $x_n$ as defined in each case.

(a) $x_n = \dfrac{1 + (-1)^n}{n}$.

(b) $x_n = \dfrac{1}{\sqrt{n}} \cos \dfrac{n\pi}{2}$.

(c) $x_n = \dfrac{2n^3 - n}{n^4 + n}$.

(d) $x_n = \dfrac{5n^3 + 2n^2 - n^3(\frac{1}{2})^n}{3n^3 + n^2(\frac{3}{4})^n}$.

(e) $x_n = \left(100 + \dfrac{1}{n}\right)^2 \cdot \left(1 + \dfrac{n-1}{n^2}\right)^{100}$.

(f) $x_n = \dfrac{\sqrt{n+1}}{\sqrt{2n+1}}$.

**2.** Which of the following sequences are not convergent? For those which are convergent, state what the limit is. The given expression is $x_n$.

(a) $n[1 + (-1)^n]$.

(b) $\dfrac{\sin n}{n}$.

(c) $n^2 + (-1)^n 2n$.

(d) $\dfrac{1}{n + 1 + (-1)^n(1 - n)}$.

(e) $\dfrac{n!}{2\cdot 4\cdot 6\cdots(2n)}$.

(f) $\dfrac{(n+1)!}{1\cdot 3\cdots(2n+1)}$.

**3.** Show that the following sequences are convergent without finding their limits.

(a) $x_n = \dfrac{n(n+1)}{2^n}$.

(b) $x_n = \dfrac{2^{2n}(n!)^2}{(2n+1)!}$.

(c) $x_n = 1 - \dfrac{5^n}{n!}$.

(d) $x_n = \dfrac{1}{n}\left[\dfrac{2\cdot 4\cdots(2n)}{1\cdot 3\cdots(2n-1)}\right]^2$.

**4.** If $x_1 = 1$, $x_2 = 3$, and generally $x_{n+1} = \frac{1}{2}(x_n + x_{n-1})$ when $n \geq 2$, use Theorem 14-C to show that $\{x_n\}$ is convergent.

**5.** Give a proof that if $\{x_n\}$ is convergent, the condition stated in Theorem 14-C is satisfied. Make use of (2) in § 14-2, with $a = x_n - x$, $b = x - x_m$.

## 14-5 L'Hospital's Rule

In this section we shall learn methods for finding the limit of a quotient of two functions in circumstances such that the limit cannot be found directly by the rule that the limit of a quotient is the quotient of the limits. This rule is Theorem 1-E of § 1-8. Suppose the problem is to find

$$\lim_{t \to a} \frac{f(t)}{g(t)}, \tag{1}$$

and suppose we know that the numerator and denominator each approaches a limit as $t \to a$:

$$\lim_{t \to a} f(t) = A, \qquad \lim_{t \to a} g(t) = B. \tag{2}$$

Then the theorem asserts that, provided $B \neq 0$, the limit in (1) has the value $A/B$. This theorem gives no information about the limit (1) if $B = 0$; it also fails to give any information if the limits in (2) do not both exist. Thus we get no information about

$$\lim_{t\to 0} \frac{e^{2t} - \cos t}{t}. \tag{3}$$

Here we let $f(t) = e^{2t} - \cos t$, $g(t) = t$; both $f(t)$ and $g(t)$ approach 0 as $t \to 0$. Yet the limit in (3) exists and is equal to 2, as we shall see presently (Example 1). Another type of situation is illustrated by the limit

$$\lim_{t\to\infty} \frac{\log t}{t}. \tag{4}$$

Here, with $f(t) = \log t$, $g(t) = t$, we cannot use the aforementioned theorem, because $f(t)$ and $g(t)$ both become infinite as $t \to \infty$. Yet the limit of the quotient exists and has the value 0, as we shall show (Example 2).

There is a systematic method for finding limits such as those in (3) and (4), provided the functions $f(t)$ and $g(t)$ meet certain requirements. This systematic method is known as l'Hospital's rule. It is named after a French mathematician who popularized it in a textbook published in 1696. The method is to consider the quotient

$$\frac{f'(t)}{g'(t)} \quad \text{instead of} \quad \frac{f(t)}{g(t)}.$$

L'Hospital's rule states that, *under certain conditions*, the second quotient has the same limit as the first; that is

$$\lim_{t\to a} \frac{f(t)}{g(t)} = \lim_{t\to a} \frac{f'(t)}{g'(t)}. \tag{5}$$

In stating the conditions under which (5) is valid, we assume that $t$ ranges over some interval having the point $t = a$ at one end, e.g., $0 < t < 1$ with $a = 0$ or $a = 1$. We assume that both functions are differentiable on this interval, and that neither $g(t)$ nor $g'(t)$ is ever equal to 0 on the interval. There are then two cases considered in l'Hospital's rule:

*Case 1.* $f(t) \to 0$ *and* $g(t) \to 0$ *as* $t \to a$.

*Case 2.* $g(t) \to +\infty$ *or* $g(t) \to -\infty$ *as* $t \to a$.

We now state the rule.

L'Hospital's Rule. *Under the conditions of either Case 1 or Case 2 and the other conditions already stated,* $f(t)/g(t)$ *has the same limit as* $f'(t)/g'(t)$, *provided the latter quotient either approaches a finite limit, or tends definitely to* $+\infty$ *or to* $-\infty$ *as* $t \to a$. *The rule is also valid with* $t \to +\infty$ *or* $t \to -\infty$ *in place of* $t \to a$, *both in hypotheses and conclusion.*

We postpone discussion of proof of the rule until after it has been illustrated by examples.

**Example 1:** The limit in (3) comes under Case 1. Hence

$$\lim_{t\to 0} \frac{e^{2t} - \cos t}{t} = \lim_{t\to 0} \frac{2e^{2t} + \sin t}{1} = 2.$$

**Example 2:** The limit in (4) comes under Case 2. Hence

$$\lim_{t\to\infty} \frac{\log t}{t} = \lim_{t\to\infty} \frac{1/t}{1} = 0.$$

**Example 3:** Find $\lim\limits_{x\to+\infty} \dfrac{e^x}{x^2}$. This comes under Case 2, with $f(x) = e^x$, $g(x) = x^2$. It does not matter, of course, that the variable is $x$ instead of $t$. Applying the rule, we have

$$\lim_{x\to+\infty} \frac{e^x}{x^2} = \lim_{x\to+\infty} \frac{e^x}{2x}.$$

The limit of the new quotient also comes under Case 2, so we apply the rule a second time.

$$\lim_{x\to+\infty} \frac{e^x}{2x} = \lim_{x\to+\infty} \frac{e^x}{2} = +\infty.$$

Thus we conclude that $e^x/x^2 \to +\infty$ as $x \to +\infty$. This implies, of course, that $e^x$ is much larger than $x^2$ when $x$ is very large.

It may be necessary to apply the rule more than twice. Also, before reapplying the rule at any stage, it may be possible to make some simplifications, such as cancellation of common factors of the numerator and denominator, or simplification by the use of trigonometric identities.

Some limit problems do not directly appear to be of the types considered in l'Hospital's rule, but may be handled by the rule after some simple preliminary work.

**Example 4:** Find $\lim\limits_{x\to 0} x \log x$. The value of this limit is not at once apparent, because $\log x \to -\infty$ as $x \to 0$, and we cannot tell by inspection whether the product $x \log x$ is more influenced by the smallness of $x$ or the large magnitude of $\log x$. To use l'Hospital's rule we write

$$x \log x = \frac{\log x}{1/x},$$

and take $f(x) = \log x$, $g(x) = 1/x$, so that we have an instance of Case 2 of the rule. Thus

$$\lim_{x\to 0} x \log x = \lim_{x\to 0} \frac{1/x}{-1/x^2} = \lim_{x\to 0} (-x) = 0.$$

**Example 5:** Find $\lim\limits_{x\to 0} (1 + \sin x)^{1/x}$. In problems where both the base and the exponent are variable, it is usually best to begin by considering the logarithm of the expression. Let

$$y = (1 + \sin x)^{1/x}, \qquad \log y = \frac{\log(1 + \sin x)}{x}.$$

As $x \to 0$ in the expression for $\log y$, we have a situation covered by Case 1 of l'Hospital's rule. Hence

$$\lim_{x \to 0} (\log y) = \lim_{x \to 0} \frac{\dfrac{\cos x}{1 + \sin x}}{1} = 1.$$

But $\log y \to 1$ implies $y \to e$. Therefore the required limit has the value $e$.

For the proof of l'Hospital's rule we need a theorem that is an extension of the law of the mean.

EXTENDED LAW OF THE MEAN. *Let the functions $F(x)$, $G(x)$ be continuous, $a \leqq x \leqq b$, and suppose that they are differentiable for values of $x$ between $a$ and $b$: $a < x < b$. Finally, suppose that $G(b) \neq G(a)$, and that the derivatives $F'(x)$, $G'(x)$ are never zero simultaneously. Then for a suitable value $x = X$, $a < X < b$, we have the formula*

$$\frac{F(b) - F(a)}{G(b) - G(a)} = \frac{F'(X)}{G'(X)}. \tag{6}$$

The proof is based on the ordinary law of the mean (Theorem 2-C). We construct the function

$$\phi(x) = F(x)[G(b) - G(a)] - G(x)[F(b) - F(a)].$$

We observe that $\phi(b) = \phi(a)$, and hence, when the law of the mean is applied, we obtain

$$0 = \phi(b) - \phi(a) = (b - a)\phi'(X),$$

or $\phi'(X) = 0$ for some $X$ between $a$ and $b$. Now

$$\phi'(x) = F'(x)[G(b) - G(a)] - G'(x)[F(b) - F(a)].$$

The equation $\phi'(X) = 0$ thus becomes exactly equation (6), the division being permissible because of our assumptions.

If we take $G(x) = x$, then (6) can be written

$$F(b) - F(a) = (b - a)F'(X);$$

this is the ordinary law of the mean.

We now turn our attention to the proof of l'Hospital's rule. There are two cases to consider:

*Case 1.* $f(t)$ *and* $g(t) \to 0$ *as* $t \to a$;

*Case 2.* $|g(t)| \to \infty$ *as* $t \to a$.

In both cases it is to be understood that $t$ approaches $a$ from one side only. The problem is to show that, if the quotient $f'(t)/g'(t)$ approaches a limit as $t \to a$, then the quotient $f(t)/g(t)$ approaches the same limit as the first mentioned quotient. It is quite easy to prove this in Case 1, provided the symbol $a$ does not denote $+\infty$ or $-\infty$. For, if we agree to define $f(a) = g(a) = 0$, both $f(t)$ and $g(t)$ will be continuous at $a$, and we can apply the extended law of the mean to these functions on an interval with $t = a$ at one end. For a point $t \neq a$ of this interval the extended law of the mean tells us that there is some $T$ between $t$ and $a$ such that

$$\frac{f(t) - f(a)}{g(t) - g(a)} = \frac{f(t)}{g(t)} = \frac{f'(T)}{g'(T)}. \tag{7}$$

If $t \to a$, then $T \to a$ also, and the quotient on the right in (7) approaches the limit denoted by $\lim\limits_{t\to a} \dfrac{f'(t)}{g'(t)}$. Hence, by (7),

$$\lim_{t\to a} \frac{f(t)}{g(t)} = \lim_{t\to a} \frac{f'(t)}{g'(t)}. \tag{8}$$

This is the conclusion we desired to reach.

The foregoing argument does not apply to Case 2, nor does it apply to Case 1 if $a = \pm\infty$. We shall indicate how to modify the argument so as to handle Case 2. The further discussion of Case 1 is left to the student in Exercise 11, with suggestions of how to proceed.

For the sake of definiteness, let us suppose that $t$ is to approach $a$ from the left. Let $s$ and $t$ be distinct points on the left of $a$, in the interval where the two functions satisfy the conditions stated in connection with l'Hospital's rule. Then there is a point $T$ between $s$ and $t$ such that

$$\frac{f(t) - f(s)}{g(t) - g(s)} = \frac{f'(T)}{g'(T)}. \tag{9}$$

This is a direct application of the extended law of the mean. If we divide numerator and denominator on the left in (9) by $g(t)$, and rearrange slightly, we obtain

$$\frac{f(t)}{g(t)} = \left[1 - \frac{g(s)}{g(t)}\right]\frac{f'(T)}{g'(T)} + \frac{f(s)}{g(t)}. \tag{10}$$

Now suppose that $s < t$, and hence $s < T < t$. We are going to outline a method of using (10) to obtain a proof of (8). The idea of the proof is this: Denote $\lim\limits_{t\to a} \dfrac{f'(t)}{g'(t)}$ by $A$. Choose $s$ so close to $a$ that $f'(u)/g'(u)$ is very near $A$ if $s < u < a$. Then $f'(T)/g'(T)$ is very near $A$, since $s < T < t$. Now, keeping $s$ fixed, let $t \to a$. Then

$$\frac{g(s)}{g(t)} \to 0 \quad \text{and} \quad \frac{f(s)}{g(t)} \to 0,$$

since the numerators here are fixed, and $|g(t)| \to \infty$ (the Case 2 hypothesis). Thus, as $t \to a$, the right side of (10) becomes about the same as $f'(T)/g'(T)$, and is therefore near $A$. Since this "nearness to $A$" can be controlled to any extent we desire by the choice of $s$ before letting $t \to a$, the conclusion from (10) is that $\lim_{t \to a} \dfrac{f(t)}{g(t)} = A$ also. In other words, (8) holds. This outline of a proof can be made more formal and precise by bringing in the exact definitions of limits in terms of inequalities.

## EXERCISES

**1.** Find each of the following limits.

(a) $\lim_{x \to 1/2} \dfrac{\log 2x}{2x - 1}$.

(b) $\lim_{x \to 0} \dfrac{e^{2x} - \cos x}{\tan 3x}$.

(c) $\lim_{x \to 0} \dfrac{\tan x - \sin x}{x^2}$.

(d) $\lim_{x \to 1} \dfrac{5x^8 - 11x^7 + 6x^6 + x^2 - x}{(x - 1)^3}$.

(e) $\lim_{x \to +\infty} \dfrac{\log(1 + e^{3x})}{x}$.

(f) $\lim_{x \to 0} \dfrac{\tan 2x - 2x}{x^3}$.

**2.** Find each of the following limits.

(a) $\lim_{x \to 2\pi} \dfrac{\sin^2(x/2)}{1 + \cos(x/2)}$.

(b) $\lim_{x \to +\infty} \dfrac{\log(4 + e^{-2x})}{e^{-2x}}$.

(c) $\lim_{x \to 0} \dfrac{\sin(\pi \cos \pi x)}{x \sin x}$.

(d) $\lim_{x \to +\infty} \dfrac{x^2 \log x}{e^x}$.

(e) $\lim_{x \to 0} \dfrac{4^x - 2^x}{x}$.

(f) $\lim_{x \to +\infty} \dfrac{(1 + x^3)^{1/2}}{\log x}$.

**3.** Prove by mathematical induction:

(a) $\lim_{x \to +\infty} \dfrac{(\log x)^n}{x} = 0, \quad n = 1, 2, \cdots$.

(b) $\lim_{x \to 0} x (\log x)^n = 0, \quad n = 1, 2, \cdots \quad (x > 0)$.

Can you deduce (b) from (a)?

(c) $\lim_{x \to +\infty} \dfrac{x^n}{e^x} = 0, \quad n = 1, 2, \cdots$.

**4.** Find $\lim_{x \to 0} x^p \log x$ if $0 < p$ (assuming $x > 0$, of course).

**5.** If $a > 1$ and $p > 0$, show that $\lim_{x \to +\infty} \dfrac{x^p}{\log_a x} = +\infty$. As a result, certainly $\log_a x < x^p$ if $x$ is sufficiently large.

**6.** Assume $a > 1$, $p > 0$, and prove that $\lim_{x \to +\infty} \frac{a^x}{x^p} = +\infty$. Let $n$ be the integer such that $n - 1 < p \leq n$ and show that the proof is achieved by $n$ applications of l'Hospital's rule. As a result, certainly $x^p < a^x$ if $x$ is sufficiently large.

**7.** Find each of the following limits.

(a) $\lim_{x \to 1} x^{1/(1-x)}$.

(b) $\lim_{x \to 0} x^x \ (x > 0)$.

(c) $\lim_{x \to +\infty} x(e^{1/x} - 1)$.

(d) $\lim_{x \to +\infty} x\,(\tan^{-1} x - \pi/2)$.

(e) $\lim_{x \to 0} (e^x + x)^{1/x}$.

(f) $\lim_{x \to +\infty} \left(\cos \frac{2}{x}\right)^{x^2}$.

**8.** Show that $\lim_{x \to 0} x^{-n}e^{-1/x^2} = 0$ for all values of $n$. Hence show that $e^{-1/x^2}$ and all its derivatives approach 0 as $x \to 0$.

**9.** Find the limits approached by $xe^{-1/x}$ as $x$ approaches 0 through negative and positive values, respectively, of $x$. Suggestion: Let $t = 1/x$.

**10.** If $y = x/(1 + e^{1/x})$, find the limits approached by $y$ and $dy/dx$ as $x$ approaches 0 (a) through positive values; (b) through negative values. *Suggestion*: Let $t = 1/x$.

**11.** Assuming that l'Hospital's rule has already been proved for the case $t \to a$, where $a$ is finite, there is a simple device for proving the rule when the limits are all taken as $t \to +\infty$. Make the change of variable $x = 1/t$ and define new functions $f_1(x) = f(t)$, $g_1(x) = g(t)$. Then the new functions satisfy the conditions of l'Hospital's rule as $x \to 0^+$ (which corresponds to $t \to +\infty$). Show that

$$\lim_{t \to +\infty} \frac{f(t)}{g(t)} = \lim_{t \to +\infty} \frac{f'(t)}{g'(t)}$$

is a consequence of the fact that

$$\lim_{x \to 0^+} \frac{f_1(x)}{g_1(x)} = \lim_{x \to 0^+} \frac{f_1'(x)}{g_1'(x)}.$$

# CHAPTER XV

# INFINITE SERIES AND TAYLOR'S FORMULA

## 15-1 Sequences and Series

For a good understanding of the subject of infinite series it is essential to know the fundamental things about sequences, as presented in § 14-4.

One particularly important and useful scheme for generating sequences is by successive additions. One of the simplest such schemes is that in which we start with some number $a$ and add to it successively the terms of a geometric progression of which $a$ is the first term. If the progression is

$$a, \quad ar, \quad ar^2, \quad \cdots, \quad ar^{n-1}, \quad \cdots,$$

the sum of the first $n$ terms is

$$S_n = a + ar + \cdots + ar^{n-1}. \tag{1}$$

This gives us a sequence $\{S_n\}$. We may obtain another formula for $S_n$ in this way: Multiply the sum in (1) by $r$ and subtract the result from the original sum. Since

$$rS_n = ar + ar^2 + \cdots + ar^n,$$

this gives us

$$(1 - r)S_n = a - ar^n, \qquad S_n = \frac{a(1 - r^n)}{1 - r}, \tag{2}$$

provided that $r \neq 1$. If $r = 1$ we have $S_n = na$, of course. Now let us ask whether the sequence $\{S_n\}$ is convergent or not. We shall suppose $a \neq 0$. There are three cases, according as $|r| < 1$, $|r| = 1$, or $|r| > 1$. If $|r| < 1$, we know from Example 4 in § 14-4 that $|r|^n \to 0$, which is the same as $r^n \to 0$. Hence, by Theorem 14-A as applied to (2),

$$\lim_{n\to\infty} S_n = \frac{a}{1-r} \quad \text{if} \quad |r| < 1. \tag{3}$$

If $|r| > 1$, the term $r^n$ in (2) gets very large in absolute value when $n$ is large; the sequence $\{S_n\}$ is not bounded, and hence does not converge. The cases $r = \pm 1$ are left for the exercises.

The formula (3) can be used to convert repeating nonterminating decimals into fractions.

**Example:** Consider the decimal $0.3737\cdots$. If we let $S_n$ be the *terminating* decimal obtained by cutting this off with $2n$ decimal places, we see that

$$S_n = \frac{37}{10^2} + \frac{37}{10^4} + \cdots + \frac{37}{10^{2n}}.$$

This has the form (1) with $a = 37/100$ and $r = 1/100$. Hence

$$\lim_{n\to\infty} S_n = \frac{0.37}{1 - 0.01} = \frac{37}{99}.$$

The very meaning of the nonterminating decimal is expressed by the limit as $n \to \infty$. Hence

$$0.3737\cdots = \frac{37}{99}.$$

Nonterminating decimals which do not have a repeating pattern are also to be regarded as limits of sequences. Let us consider *any* decimal

$$0.a_1a_2a_3\cdots, \tag{4}$$

where each $a_n$ denotes one of the digits $0, 1, \cdots, 9$ according to some definite rule. The rule may be very complicated, however, as for instance in the decimal for $\pi/3$, where each digit $a_n$ is definitely determined, but one cannot easily find out what $a_{1000}$ is. In the case of the decimal (4) let us define $S_n = 0.a_1a_2\cdots a_n$, stopping with $n$ decimal places. We can also write

$$S_n = \frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n}.$$

It is clear that $\{S_n\}$ is a nondecreasing sequence. Moreover, the sequence is bounded, for certainly $S_n < 1$, no matter how large $n$ is or how the digits $a_1, \cdots, a_n$ are chosen. The sequence is therefore convergent, with its least upper bound as limit, by Theorem 14-B. This limit of $S_n$ is precisely what the complete decimal (4) denotes.

From these particular cases we now pass to the general notion of a sequence produced by successive additions or subtractions. Let $u_1, u_2, u_3, \cdots$ be a sequence of numbers (they can be positive, negative, or zero) and let $\{S_n\}$ be the sequence of sums formed in this way: $S_1 = u_1$, $S_2 = u_1 + u_2$, and in general

$$S_n = u_1 + \cdots + u_n. \tag{5}$$

The study of infinite series is the study of sequences formed in this way. It is customary to write down the expression

$$u_1 + u_2 + u_3 + \cdots + u_n + \cdots \tag{6}$$

and call it the *infinite series* with *terms* $u_1, u_2, \cdots, u_n, \cdots$. This expression is to be thought of simply as an agglomeration of symbols which shows us what the terms are and which by its plus signs suggests the process of forming the sequence $S_1, S_2, S_3, \cdots$ in the manner already indicated. We call $S_n$ the $n$th *partial sum* of the series.

We can consider this expression (6) regardless of whether or not the sequence $\{S_n\}$ converges. If the *sequence* converges we call the series *convergent*. If the sequence does not converge we call the series *divergent*, or say that it diverges. If the sequence converges, with limit $S$, we call $S$ the *sum* (or value) of the series. Otherwise we do not speak of any sum for the series. When the series is convergent with sum $S$ we commonly write

$$S = u_1 + u_2 + \cdots + u_n + \cdots.$$

Our program of study in this chapter has two principal aims. For one thing, we show some of the ways in which infinite series are used in connection with calculus. We shall find, for example, that each of the elementary functions $\log(1 + x)$, $\tan^{-1} x$, $\sin x$, $e^x$, and many others, can be expressed as the sum of a certain infinite series whose terms are formed in a rather simple way. Such series representations are useful for obtaining numerical calculations. Also, the whole idea of representing functions by infinite series is a fruitful one, and it suggests a powerful method for the construction of many new functions.

The second main aim of our study of infinite series is to learn in systematic fashion some of the most important ways of testing to find out whether a given series is convergent or divergent.

One interesting aspect of the study of infinite series is that it provides many stimuli for plain curiosity and offers many results which stir the imagination. Consider, for example, the series

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} + \cdots. \tag{7}$$

As we shall see later, this series is convergent. Its sum can be shown to be $\pi^2/6$, but methods beyond the scope of this book are required. On the other hand, the series

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots \tag{8}$$

is divergent (see Exercise 6). This is called the harmonic series. If we change the signs of alternate terms, we get a convergent series whose sum, as we shall see in § 15-2, is log 2:

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots.$$

The $n$th term here is $(-1)^{n+1}/n$.

We conclude this section with a simple but important theorem.

THEOREM 15-A. *If a series is convergent, and $u_n$ is the nth term, then $u_n \to 0$ as $n \to \infty$. Or, to put the matter another way, if $u_n$ does not approach 0 as $n \to \infty$, the series cannot be convergent.*

*Proof.* From (5) we see that $u_n = S_n - S_{n-1}$. Now if $S_n \to S$, we can write

$$u_n = (S_n - S) + (S - S_{n-1}),$$

and so

$$|u_n| \leq |S_n - S| + |S - S_{n-1}|.$$

As $n$ gets large, $|S_n - S|$ and $|S - S_{n-1}|$ both approach 0, and hence so does $u_n$.

The student must not confuse Theorem 15-A with the converse proposition, which is not true. One cannot conclude that a series is convergent merely because $u_n \to 0$. The example of the series (8) shows this. Even though $1/n \to 0$, it can be shown that the sum of the first $2^n$ terms of the series (8) is not less than $(n + 2)/2$, and hence the sequence of partial sums has no upper bound.

## EXERCISES

**1.** Find the fraction represented by the indicated repeating nonterminating decimals.

(a) $0.444\cdots$.
(b) $0.132132\cdots$.
(c) $3.1454545\cdots$.
(d) $2.999\cdots$.

**2.** Find the sum of each series.

(a) $1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \cdots + \frac{1}{3^{n-1}} + \cdots$.

(b) $4 + 2 + 1 + \cdots + \frac{1}{2^{n-3}} + \cdots$.

**3.** Which of these series are convergent and which are divergent?

(a) $1 - 1 + 1 - 1 + \cdots + (-1)^{n+1} + \cdots$.
(b) $\sin \pi + \sin 2\pi + \cdots + \sin n\pi + \cdots$.
(c) $\sin (\pi/2) + \sin (2\pi/2) + \cdots + \sin (n\pi/2) + \cdots$.
(d) $\log_{10} 5^{1/2} + \log_{10} 5^{1/4} + \cdots + \log_{10} 5^{1/2^n} + \cdots$.

(e) $\frac{1+1}{100+2} + \frac{1+2}{100+4} + \cdots + \frac{1+n}{100+2n} + \cdots$.

(f) $\frac{1}{\sqrt{1+100}} + \frac{2}{\sqrt{1+400}} + \cdots + \frac{n}{\sqrt{1+100n^2}} + \cdots$.

4. Discuss the convergence or divergence of the series

$$a + ar + ar^2 + \cdots + ar^{n-1} + \cdots$$

if $a \neq 0$ and $r = \pm 1$.

5. Suppose that we have two convergent series, one with $n$th term $a_n$ and sum $A$, the other with $n$th term $b_n$ and sum $B$. What can you say about the infinite series with $n$th term $c_n$, where $c_n = a_n + b_n$? Give your reasons, and cite a theorem for support.

6. Let $\{S_n\}$ be the sequence of partial sums of the harmonic series (8). Consider $S_2$, $S_4$, $S_8$, $S_{16}$ and so on, and show that each one of these exceeds its predecessor by more than $\frac{1}{2}$. Thus, by induction, demonstrate that $S_{2^n} > (n + 2)/2$ if $n > 1$.

## 15-2 Various Series Derived from Geometric Progressions

The series with terms $1, t, t^2, \cdots, t^n, \cdots$ is convergent if $|t| < 1$, for the sum of the first $n$ terms is

$$s_n = 1 + t + \cdots + t^{n-1} = \frac{1 - t^n}{1 - t}, \tag{1}$$

by the results on geometric progressions in § 15-1. Since $|t| < 1$, the sequence $\{s_n\}$ is convergent, with limit $(1 - t)^{-1}$, and we obtain

$$\frac{1}{1 - t} = 1 + t + t^2 + \cdots + t^{n-1} + \cdots \quad \text{if} \quad |t| < 1. \tag{2}$$

The series here is called a *geometric series.*

We can get other series from (2) by substituting various things in place of $t$. For example, we can replace $t$ by $-x^2$, provided that $x^2 < 1$, for we must have $|t| < 1$. The result is

$$\frac{1}{1 + x^2} = 1 - x^2 + x^4 - \cdots + (-1)^{n-1}x^{2n-2} + \cdots.$$

An even greater variety of results can be obtained by working with (1) in various ways, as we shall now show. We rewrite (1) in the form

$$\frac{1}{1 - t} = 1 + t + \cdots + t^{n-1} + \frac{t^n}{1 - t}. \tag{3}$$

Now choose any number $x$ such that $-1 \leq x < 1$ and integrate both sides of equation (3) from 0 to $x$. The result is

$$-\log (1 - x) = x + \frac{x^2}{2} + \cdots + \frac{x^n}{n} + \int_0^x \frac{t^n}{1 - t}\, dt.$$

From this will follow the infinite series formula

$$-\log (1 - x) = x + \frac{x^2}{2} + \cdots + \frac{x^n}{n} + \cdots \tag{4}$$

as soon as we prove that

$$\lim_{n\to\infty} \int_0^x \frac{t^n}{1-t}\,dt = 0. \tag{5}$$

Putting $x = -1$ in (4) gives a result that was referred to in § 15-1. For convenience we denote the value of the integral in (5) by $I_n(x)$. We proceed to get some estimates to show that $I_n(x)$ is small when $n$ is large. If $0 < x < 1$, then

$$0 \le \frac{t^n}{1-t} \le \frac{t^n}{1-x} \quad \text{when} \quad 0 \le t \le x,$$

and so in this case,

$$0 < \int_0^x \frac{t^n}{1-t}\,dt \le \int_0^x \frac{t^n}{1-x}\,dt = \frac{x^{n+1}}{(n+1)(1-x)}. \tag{6}$$

Hence certainly $I_n(x) \to 0$ if $0 < x < 1$. If $-1 \le x < 0$ it is not difficult to show that

$$|I_n(x)| \le \frac{|x|^{n+1}}{n+1}, \tag{7}$$

and so $I_n(x) \to 0$ in this case also. The proof of (7) is left as an exercise.

Further discussion of (4) will be taken up presently. First, however, we indicate how another interesting formula can be deduced from (3). This time we put $-t^2$ in place of $t$, getting

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - \cdots + (-1)^{n-1}t^{2n-2} + (-1)^n \frac{t^{2n}}{1+t^2}.$$

Choosing $x$ such that $|x| \le 1$, we integrate this formula from 0 to $x$, getting

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^{n-1}\frac{x^{2n-1}}{2n-1} + I_n(x),$$

where
$$I_n(x) = (-1)^n \int_0^x \frac{t^{2n}}{1+t^2}\,dt. \tag{8}$$

By estimating the size of this integral it is not hard to show that $I_n(x) \to 0$ as $n \to \infty$. In this way we derive the series formula

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^{n-1}\frac{x^{2n-1}}{2n-1} + \cdots, \tag{9}$$

valid if $-1 \le x \le 1$. A particular case is that in which $x = 1$:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots + (-1)^{n-1}\frac{1}{2n-1} + \cdots. \tag{10}$$

There are many devices, some of them quite intricate, for obtaining formulas in which (9) may be used so as to yield an effective method of computing $\pi$. Formula (10), though interesting, is not of much use for computation.

**Example 1:** We can show that

$$\frac{\pi}{4} = \tan^{-1}\frac{1}{2} + \tan^{-1}\frac{1}{3} \tag{11}$$

and use this result. Suppose $\alpha = \tan^{-1}\frac{1}{2}$, $\beta = \tan^{-1}\frac{1}{3}$. Then

$$\tan(\alpha + \beta) = \frac{\tan\alpha + \tan\beta}{1 - \tan\alpha\tan\beta} = \frac{\frac{1}{2} + \frac{1}{3}}{1 - \frac{1}{6}} = 1.$$

Therefore $\alpha + \beta = \pi/4$. This proves (11). Now, by (9),

$$\tan^{-1}\frac{1}{2} = \frac{1}{2} - \frac{1}{3}\left(\frac{1}{2}\right)^3 + \frac{1}{5}\left(\frac{1}{2}\right)^5 - \cdots.$$

$$\tan^{-1}\frac{1}{3} = \frac{1}{3} - \frac{1}{3}\left(\frac{1}{3}\right)^3 + \frac{1}{5}\left(\frac{1}{3}\right)^5 - \cdots.$$

If we take $n$ terms of each series, we get an approximation to $\pi/4$. For $n = 1$, 2, 3, 4, 5 these approximations are, respectively,

$$n = 1: \quad \frac{1}{2} + \frac{1}{3} = 0.83333\cdots.$$

$$n = 2: \quad 0.8333 - \frac{1}{3}\left(\frac{1}{2^3} + \frac{1}{3^3}\right) = 0.7793.$$

$$n = 3: \quad 0.7793 + \frac{1}{5}\left(\frac{1}{2^5} + \frac{1}{3^5}\right) = 0.7864.$$

$$n = 4: \quad 0.7864 - \frac{1}{7}\left(\frac{1}{2^7} + \frac{1}{3^7}\right) = 0.7852.$$

$$n = 5: \quad 0.7852 + \frac{1}{9}\left(\frac{1}{2^9} + \frac{1}{3^9}\right) = 0.7854.$$

Thus, approximately

$$\frac{\pi}{4} = 0.7854.$$

Let us now return to the series (4). We can use it to obtain a series formula for $\log y$, where $y$ is any positive number. First we observe the result of replacing $x$ by $-x$ in (4):

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + \cdots. \tag{12}$$

On combining (4) and (12) by addition (see Exercise 5 in § 15-1) we obtain

$$\frac{1}{2}\log\frac{1 + x}{1 - x} = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{2n-1}}{2n - 1} + \cdots. \tag{13}$$

Now, if $y > 0$, let

$$x = \frac{y - 1}{y + 1}, \quad \text{so that} \quad y = \frac{1 + x}{1 - x}.$$

Then $-1 < x < 1$, and so

$$\log y = \log \frac{1+x}{1-x} = 2\left(x + \frac{1}{3}x^3 + \frac{1}{5}x^5 + \cdots\right). \tag{14}$$

If $y$ is large, $x$ is near 1, and this series does not converge rapidly enough to be very useful for numerical computation. But if $y$ is near 1, $x$ is near 0, and the series converges very rapidly if $x$ is quite small.

**Example 2:** If $x = \frac{1}{10}$, $y = \frac{11}{9} = 1.222\cdots$. By (14) we have

$$\log \frac{11}{9} = 2\left[\frac{1}{10} + \frac{1}{3000} + \frac{1}{500{,}000} + \cdots\right].$$

Using just three terms of the series we get

$$\log \frac{11}{9} = 0.20067,$$

a result that is accurate to five decimal places, because the latter terms of the series are all too small to affect the first five decimal places.

## EXERCISES

**1.** (a) From (8) show that $|I_n(x)| \le \dfrac{|x|^{2n+1}}{2n+1}$, and hence that $I_n(x) \to 0$ if $|x| \le 1$. Notice that the estimate of $I_n(x)$ means that the sum of the first $n$ terms of (9) differs from $\tan^{-1} x$ by not more than the absolute value of the first term not taken. (b) On this basis how many terms would be needed in (10) to get an approximation of $\pi/4$ accurate to 4 decimal places? (c) Show that the five-term approximation of $\pi/4$ in Example 2 is too large, but not by as much as $5\cdot 10^{-5}$. This does not allow for round-off error.

**2.** (a) Deduce the series

$$\log(1+y) = \log y + 2\left[\frac{1}{2y+1} + \frac{1}{3}\left(\frac{1}{2y+1}\right)^3 + \frac{1}{5}\left(\frac{1}{2y+1}\right)^5 + \cdots\right]$$

by putting $x = (2y+1)^{-1}$ in (13).
(b) Given $\log 5 = 1.60944$, $\log 10 = 2.30259$, and $\log 20 = 2.99573$, find $\log 6$, $\log 11$, and $\log 21$.

**3.** (a) Show that $\pi/4 = 4\tan^{-1}\frac{1}{5} - \tan^{-1}\frac{1}{239}$ by setting $\alpha = \tan^{-1}\frac{1}{5}$, $\beta = \tan^{-1}\frac{1}{239}$ and computing successively $\tan 2\alpha$, $\tan 4\alpha$, $\tan(4\alpha - \beta)$.
(b) Use the result in (a) along with series (9) to compute $\pi$, using five terms of the series for $\alpha$ and two terms of the series for $\beta$.

**4.** Prove the inequality (7) under the stated conditions.

**5.** Derive a formula

$$\frac{1}{2}\log\frac{1+x}{1-x} = x + \frac{x^3}{3} + \cdots + \frac{x^{2n-1}}{2n-1} + I_n(x),$$

where $I_n(x)$ is a certain integral, by putting $t^2$ in place of $t$ in (3), and integrating. Show that

$$|I_n(x)| \leq \frac{|x|^{2n+1}}{(2n+1)(1-x^2)} \quad \text{if} \quad x^2 < 1.$$

6. Write formula (3) with $n + 1$ in place of $n$. Then differentiate both sides to obtain

$$\frac{1}{(1-t)^2} = 1 + 2t + 3t^2 + \cdots + nt^{n-1} + D_n,$$

where $$D_n = t^n \frac{n+1-nt}{(1-t)^2}.$$

This gives an infinite series for $(1-t)^{-2}$ if it can be proved that $D_n \to 0$. Prove this if $|t| < 1$. Observe that this will follow if it can be proved that $(2n+1)r^n \to 0$ for an $r$ such that $0 < r < 1$. Devise such a proof by refining the argument in Example 4, § 14-4, with the binomial expansion of $(1+h)^n$ carried through one more term.

## 15-3 Taylor's Formula with Integral Remainder

The series formulas of § 15-2 were obtained by a special method, starting in each case from a formula based on a geometric progression. If we wish to find series formulas for other functions, geometric progressions will not help us in most cases. We shall now consider a method of much greater generality than the method used in § 15-2.

Suppose $f(x)$ is a function which has continuous derivatives of orders $1, 2, \cdots, n+1$. We know from Chapter VI that

$$f(x) - f(a) = \int_a^x f'(t)\,dt. \tag{1}$$

Let us now integrate by parts, setting

$$u = f'(t), \qquad du = f''(t)\,dt,$$
$$dv = dt, \qquad v = -(x-t).$$

Then $$\int_a^x f'(t)\,dt = -f'(t)(x-t)\Big|_{t=a}^{t=x} + \int_a^x f''(t)(x-t)\,dt,$$

and so (1) can be written

$$f(x) = f(a) + f'(a)(x-a) + \int_a^x f''(t)(x-t)\,dt.$$

Again we use integration by parts, setting

$$u = f''(t), \qquad du = f^{(3)}(t)\,dt,$$
$$dv = (x-t)\,dt, \qquad v = -\tfrac{1}{2}(x-t)^2.$$

This time the result is

$$f(x) = f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2 + \tfrac{1}{2}\int_a^x f^{(3)}(t)(x-t)^2\,dt.$$

The process can be repeated. The general integration by parts formula is

$$\int_a^x f^{(p)}(t)(x-t)^{p-1}\,dt = \frac{1}{p}f^{(p)}(a)(x-a)^p + \frac{1}{p}\int_a^x f^{(p+1)}(t)(x-t)^p\,dt.$$

If the process is carried out until $p = n$, we get the formula

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt. \quad (2)$$

Here it is assumed that $a$ and $x$ are points of an interval on which the function $f(x)$ and its $n+1$ derivatives are continuous. This is called *Taylor's formula with integral remainder* (named after Brook Taylor, 1685–1731, an Englishman).

**Example 1:** For $f(x) = e^x$, with $a = 0$, Taylor's formula is

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \frac{1}{n!}\int_0^x e^t(x-t)^n\,dt. \quad (3)$$

Verification is left to the student.

**Example 2:** For $f(x) = \sin x$, with $a = 0$, $n = 5$, we have

$$\begin{aligned} f(x) &= \sin x, & f(0) &= 0 \\ f'(x) &= \cos x, & f'(0) &= 1 \\ f''(x) &= -\sin x, & f''(0) &= 0 \\ f^{(3)}(x) &= -\cos x, & f^{(3)}(0) &= -1 \\ f^{(4)}(x) &= \sin x, & f^{(4)}(0) &= 0 \\ f^{(5)}(x) &= \cos x, & f^{(5)}(0) &= 1 \\ f^{(6)}(x) &= -\sin x. \end{aligned}$$

Therefore

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{5!}\int_0^x (x-t)^5 \sin t\,dt.$$

Now let us suppose that $f(x)$ has continuous derivatives of all orders on some interval including the point $x = a$. In Taylor's formula let us set

$$R_n(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt, \quad (4)$$

so that, by (2),

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n + R_n(x). \quad (5)$$

From this we see that the infinite series formula

$$f(x) = f(a) + f'(a)(x - a) + \cdots + \frac{1}{n!}f^{(n)}(a)(x - a)^n + \cdots \tag{6}$$

will be valid provided that

$$\lim_{n\to\infty} R_n(x) = 0. \tag{7}$$

The series formula (6) is called the *Taylor's expansion of* $f(x)$ *in powers of* $(x - a)$. We refer to $R_n(x)$ as the *remainder* in Taylor's series. It is the error that is committed if we stop with the $n$th power of $(x - a)$ in Taylor's series. Formula (4) expresses $R_n(x)$ as a definite integral. In § 15-4 we shall find other formulas for the remainder.

Sometimes we can prove that (7) holds by estimating the size of $R_n(x)$ from the integral formula (4).

**Example 3:** Prove the validity of the series formula

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots \tag{8}$$

by showing that (7) holds for the remainder term in (3).

If $x > 0$ and $0 \le t \le x$, we know that $e^t \le e^x$. Therefore

$$0 < \frac{1}{n!}\int_0^x e^t(x - t)^n\,dt \le \frac{e^x}{n!}\int_0^x (x - t)^n\,dt = \frac{x^{n+1}e^x}{(n + 1)!}.$$

If $x < 0$ and $x \le t \le 0$, we use the inequality $e^t \le 1$, and in this case we obtain

$$\left|\frac{1}{n!}\int_0^x e^t(x - t)^n\,dt\right| \le \frac{1}{n!}\int_x^0 (t - x)^n\,dt = \frac{(-x)^{n+1}}{(n + 1)!}.$$

We now have estimates for the size of $R_n(x)$. They can be combined in the single estimate

$$\left|\frac{1}{n!}\int_0^x e^t(x - t)^n\,dt\right| \le \frac{M|x|^{n+1}}{(n + 1)!}, \tag{9}$$

where $M$ is the larger of the two numbers, 1 and $e^x$. As $n \to \infty$, the quantity on the right in (9) approaches zero, and so $R_n(x) \to 0$. This proves (8). The crucial point in this argument is the fact that, if $c$ is any positive constant, the terms of the sequence

$$c,\ \frac{c^2}{2!},\ \frac{c^3}{3!},\ \ldots,\ \frac{c^n}{n!},\ \frac{c^{n+1}}{(n + 1)!},\ \ldots$$

approach zero as $n \to \infty$, i.e.,

$$\lim_{n\to\infty} \frac{c^n}{n!} = 0. \tag{10}$$

A method of proving this is suggested in Exercise 16.

No general comprehensive rules for the validity of the Taylor's series formula are given in this book. The validity of the formula in a number of important special cases is proved in various exercises and examples, however. Lest there be a misunderstanding, we state explicitly that the Taylor's series formula (6) is not always valid. Whether or not it is valid in a particular case will depend on the nature of the particular function $f(x)$ and on the particular values of $x$ and $a$.

For the purpose of familiarizing the student with Taylor's series we shall give some examples and exercises in which the primary purpose is to obtain the Taylor's series for various functions. For this purpose the emphasis will be on the actual calculation of the successive derivatives of $f(x)$ and their evaluation at the point $x = a$. *It will be assumed without proof that in the problems of this kind which are given in this book, the Taylor's series formula* (6) *is actually true whenever the series is convergent.*

**Example 4:** The following four formulas are instances of Taylor's series with $a = 0$. They are valid for all values of $x$.

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots. \tag{11}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots. \tag{12}$$

$$\frac{1}{2}(e^x + e^{-x}) = \cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots. \tag{13}$$

$$\frac{1}{2}(e^x - e^{-x}) = \sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots. \tag{14}$$

The beginning of the series (11) was obtained in Example 2. The derivatives of $\sin x$ repeat in groups of four, so that if $f(x) = \sin x$, we get

$$\begin{aligned} f(0) &= f^{(4)}(0) = f^{(8)}(0) = \cdots = 0 \\ f'(0) &= f^{(5)}(0) = f^{(9)}(0) = \cdots = 1 \\ f''(0) &= f^{(6)}(0) = f^{(10)}(0) = \cdots = 0 \\ f^{(3)}(0) &= f^{(7)}(0) = f^{(11)}(0) = \cdots = -1. \end{aligned}$$

Thus the series (11) contains only the odd powers of $x$, and the signs on these terms are alternately plus and minus. The general term of (11) can be displayed as

$$(-1)^{n+1}\frac{x^{2n-1}}{(2n-1)!}, \qquad n = 1, 2, 3, \cdots. \tag{15}$$

Discussion of the series (12), (13) and (14) is left for the Exercises.

In Taylor's series (6) the "general" term is

$$\frac{1}{n!}f^{(n)}(a)(x-a)^n.$$

This formula gives the initial term $f(a)$ provided we follow the conventions that $0! = 1$ and $f^{(0)}(a) = f(a)$. These are standard conventions, and we shall adhere to them.

The special case of Taylor's series in which $a = 0$ is often called *Maclaurin's series.*

In order to find a general formula for the terms in the Taylor's series of a function, it is necessary to arrange the results of successive differentiations with care, so that, if possible, the general law may be discerned. This is not always easy, for the differentiations often become more and more complicated.

**Example 5:** Find Maclaurin's series for $(1 - x)^{-1/2}$.

$$f(x) = (1 - x)^{-1/2}, \qquad f(0) = 1,$$

$$f'(x) = \frac{1}{2}(1 - x)^{-3/2}, \qquad f'(0) = \frac{1}{2},$$

$$f''(x) = \frac{1\cdot 3}{2^2}(1 - x)^{-5/2}, \qquad f''(0) = \frac{1\cdot 3}{2^2},$$

$$f'''(x) = \frac{1\cdot 3\cdot 5}{2^3}(1 - x)^{-7/2}, \qquad f'''(0) = \frac{1\cdot 3\cdot 5}{2^3},$$

We need not carry the computation further, for the law of formation is now evident. The coefficients have the form

$$\frac{f^{(n)}(0)}{n!} = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n n!} = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots(2n)}.$$

The series is

$$(1 - x)^{-1/2} = 1 + \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 + \cdots. \tag{16}$$

It can be shown that this series converges if $-1 \leq x < 1$. See Exercise 6 in § 15-9.

## EXERCISES

1. Derive (12) and give a formula comparable to (15) for the general term of (12).

2. (a) Derive (13) and (14), using (6). Give formulas for the general term in each case. (b) Find the series comparable to (8) for $e^{-x}$, and by combining it with (8), give alternative derivations of (13) and (14).

3. Show that Taylor's series for $\log x$ in powers of $x - a$ (where $a > 0$) is

$$\log x = \log a + \left(\frac{x-a}{a}\right) - \frac{1}{2}\left(\frac{x-a}{a}\right)^2 + \cdots + (-1)^{n+1}\frac{1}{n}\left(\frac{x-a}{a}\right)^n + \cdots.$$

4. Calculate each of the following Maclaurin series, and supply a formula for the general term in each case. In some instances the general formula may not be applicable to the first few terms.

(a) $(1+x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{2\cdot 4}x^2 + \frac{1\cdot 3}{2\cdot 4\cdot 6}x^3 + \cdots;$

(b) $(1+x)^{3/2} = 1 + \frac{3}{2}x + \frac{3\cdot 1}{2\cdot 4}x^2 + \frac{3\cdot 1\cdot(-1)}{2\cdot 4\cdot 6}x^3 + \cdots;$

(c) $(1-x)^{-2} = 1 + 2x + 3x^2 + 4x^3 + \cdots;$

(d) $(1-x)^{-3} = 1 + 3x + \frac{3\cdot 4}{1\cdot 2}x^2 + \frac{3\cdot 4\cdot 5}{1\cdot 2\cdot 3}x^3 + \cdots;$

(e) $(1-x)^{-3/2} = 1 + \frac{3}{2}x + \frac{3\cdot 5}{2\cdot 4}x^2 + \frac{3\cdot 5\cdot 7}{2\cdot 4\cdot 6}x^3 + \cdots.$

**5.** Find (a) the Taylor's series for $5x^2 - 3x + 2$ in powers of $x + 1$; (b) the Taylor's series for $x^3$ in powers of $x - 2$. (c) Check the results in (a) and (b) by algebra.

**6.** If $P(x)$ is a polynomial of degree $n$, explain why

$$P(x) = P(a) + \frac{P'(a)}{1!}(x-a) + \cdots + \frac{P^{(n)}(a)}{n!}(x-a)^n.$$

**7.** Find the Taylor's series for $\sin x$ and $\cos x$ in powers of $x - a$. Note the results in particular when $a = \pi/4$.

**8.** (a) If $f(x) = (1 + e^x)^2$, show that $f^{(n)}(0) = 2 + 2^n$, $n = 1, 2, \cdots$, and write Maclaurin's series for the function.
(b) If $f(x) = (1 + e^x)^3$, obtain a formula for $f^{(n)}(0)$, and write Maclaurin's series for the function.

In Exercises 9-15 develop each Maclaurin's series as far as indicated.

**9.** $e^{x^2-x} = 1 - x + \frac{3}{2}x^2 - \frac{7}{6}x^3 + \frac{25}{24}x^4 + \cdots.$

**10.** $(1 + e^x)^{-1} = \frac{1}{2} - \frac{1}{4}x + \frac{1}{48}x^3 + \cdots.$

**11.** $\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots.$

**12.** $\log \cos x = -\frac{x^2}{2} - \frac{x^4}{12} - \frac{x^6}{45} - \cdots.$

**13.** $\log(1 + \sin x) = x - \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots.$

**14.** $e^{\sin x} = 1 + x + \frac{1}{2}x^2 - \frac{1}{8}x^4 + \cdots.$

**15.** $\log(1 + e^x) = \log 2 + \frac{1}{2}x + \frac{1}{8}x^2 - \frac{1}{192}x^4 + \cdots.$

**16.** Suppose $c > 0$. Choose a positive integer $p$ so large that $2c \le p$. Then show that, if $n > p$,

$$\frac{c^n}{n!} < \frac{c^p}{p!}\left(\frac{1}{2}\right)^{n-p}.$$

Use this to prove (10).

**17.** Show, using (8) and (9), that the approximation

$$e = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{7!}$$

gives the value of $e$ with an error less than $(7.5)10^{-5}$. Compute each term of the approximation to six decimal places, and show that $2.71825 < e < 2.71832$.

**18.** Prove, somewhat as in Example 3, that (7) holds for $f(x) = \sin x$ and $f(x) = \cos x$. For these cases observe that the estimate corresponding to (9) says that $|R_n(x)| \leq$ absolute value of term of series involving $x^{n+1}$.

## 15-4 Derivative Forms of the Remainder

Let us look at Taylor's formula again. It has the appearance

$$f(x) = f(a) + f'(a)(x - a) + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_n(x). \quad (1)$$

In § 15-3 we showed that $R_n(x)$ could be expressed as an integral involving the derivative of $f$ of order $n + 1$. Now we shall be interested in other ways of expressing $R_n(x)$. One thing which we notice at once is this: the formula is automatically correct if we regard (1) as the *definition* of $R_n(x)$. We shall now take this point of view, so $R_n(x)$ is defined to be whatever it takes to make (1) true when $x$ and $n$ have been fixed. With this definition of $R_n(x)$ nothing remarkable has been accomplished. The real gain will come if we can prove that $R_n(x)$ is also equal to something that can be computed by some other formula. We notice, for instance, that the $I_n(x)$ defined in connection with $-\log(1 - x)$ in § 15-2 is the $R_n(x)$ of Taylor's formula for this function, with $a = 0$, even though the formula in § 15-2 was obtained in a quite different way.

In order to keep things fairly simple, let us temporarily assume that $n = 1$. With $x$ fixed let us define a constant $M$ as

$$M = R_1(x) \Big/ \frac{(x - a)^2}{2!}, \quad \text{so that} \quad R_1(x) = \frac{(x - a)^2}{2!} M. \quad (2)$$

Then (1) becomes for this case

$$f(x) = f(a) + f'(a)(x - a) + \frac{(x - a)^2}{2!} M. \quad (3)$$

In order to find out more about $M$, let us define a function

$$\phi(u) = f(x) - f(u) - f'(u)(x - u) - \frac{(x - u)^2}{2!} M. \quad (4)$$

Notice the way $\phi(u)$ is constructed, by putting all the terms of (3) on the left side and then replacing $a$ by $u$. The motivation for constructing this function is not immediately apparent. The usefulness of the device was discovered by someone long ago, and the device has become part of the knowledge of professional mathematicians. We consider $u$ the only variable at this moment. Direct substitution shows that $\phi(x) = 0$, and (3) shows that $\phi(a) = 0$. Hence, by the law of the mean (Theorem 2-B) as applied to $\phi$ on the interval with ends at $a$ and $x$, there must be some value of $u$ between $a$ and $x$ for which $\phi'(u) = 0$. Now, from (4),

$$\begin{aligned}\phi'(u) &= -f'(u) + f'(u) - f''(u)(x-u) + (x-u)M \\ &= (x-u)[M - f''(u)]. \end{aligned} \tag{5}$$

If this value of $u$ where $\phi'(u) = 0$ is denoted by $X$, we see from (5) that $M = f''(X)$. This value of $M$ is now put back in (2), and we have

$$R_1(x) = \frac{(x-a)^2}{2!} f''(X). \tag{6}$$

The number $X$ depends on $x$, but we do not know the exact way in which it does. All we know is that $X$ is between $a$ and $x$.

This procedure can be generalized so as to work for larger values of $n$. The general result is stated in the following theorem.

THEOREM 15-B. *Suppose $f$ has continuous derivatives of orders* $1, 2, \cdots, n$ *when* $\alpha \leq x \leq \beta$, *and a derivative of order* $n + 1$ *when* $\alpha < x < \beta$. *Then if $a$ and $x$ are any two different numbers on the interval* $[\alpha, \beta]$, *the remainder* $R_n(x)$ *in Taylor's formula can be put in the form*

$$R_n(x) = \frac{f^{(n+1)}(X)}{(n+1)!}(x-a)^{n+1}, \tag{7}$$

*where $X$ is between $a$ and $x$.*

The formula (7) is called *Lagrange's form of the remainder* in Taylor's formula. It is named after J. L. Lagrange (1736-1813), a native of Turin who distinguished himself successively there, in Berlin, and in Paris. For the discussion of the proof when $n > 1$ see Exercise 9. It should be noted that the case $n = 0$ of Theorem 15-B is the law of the mean (Theorem 2-B in slightly different notation).

The Lagrange formula is useful in dealing with the Taylor's series for a function of the type $f(x) = (1 + x)^m$, where $m$ is not a positive integer.

**Example 1:** Prove for $f(x) = (1 + x)^{-1/2}$ and $a = 0$ that $R_n(x) \to 0$ if $0 < x < 1$.

By calculating systematically we find the general formula

$$f^{(k)}(x) = (-1)^k \frac{1 \cdot 3 \cdots (2k-1)}{2^k}(1+x)^{-(2k+1)/2}, \qquad k \geq 1.$$

Hence in this case, taking $k = n + 1$ and using (7), we obtain

$$R_n(x) = (-1)^{n+1} \frac{1 \cdot 3 \cdots (2n+1)}{2^{n+1}(n+1)!} \frac{x^{n+1}}{(1+X)^{(2n+3)/2}}.$$

Now, we are assuming $0 < x < 1$, and therefore $0 < X < x$. Hence certainly

$$|R_n(x)| < \frac{1 \cdot 3 \cdots (2n+1)}{2^{n+1}(n+1)!} x^{n+1}.$$

We now have to prove that the expression on the right approaches 0 as $n \to \infty$. Let us write

$$c_n = \frac{1\cdot 3\cdots(2n+1)}{2^{n+1}(n+1)!}x^{n+1}.$$

If we write what $c_{n+1}$ is and compare it with $c_n$, we find that

$$c_{n+1} = \frac{2n+3}{2n+4}xc_n < xc_n.$$

We then reason as follows:

$$c_2 < xc_1, \qquad c_3 < xc_2 < x^2c_1,$$

and in general (by induction) $c_{n+1} < x^n c_1$. Since $0 < x < 1$ and the $c_n$'s are all positive it follows that $c_n \to 0$ as $n \to \infty$. But then $R_n(x) \to 0$.

There is yet another useful formula for $R_n(x)$. It is the following. We let $x - a = h$. Then there is a certain number $\theta$, depending in some way on $x$, and such that $0 < \theta < 1$ and

$$R_n(x) = \frac{h^{n+1}(1-\theta)^n}{n!}f^{(n+1)}(a+\theta h). \tag{8}$$

This formula is due to Cauchy. Observe that $a + \theta h$ is a number between $a$ and $x$, but it is not usually the same as the $X$ in Lagrange's formula. Some work on the derivation of (8) is indicated in Exercise 10.

**Example 2:** Use (8) to prove that $R_n(x) \to 0$ in the case of $f(x) = (1+x)^{-1/2}$, $a = 0$, if $-1 < x < 0$.

We utilize the calculation of the derivatives made in Example 1. From (8) we find that $R_n(x)$ can be written in the form

$$R_n(x) = (-1)^{n+1}\frac{1\cdot 3\cdots(2n+1)}{n!2^{n+1}}\left(\frac{1-\theta}{1+\theta x}\right)^n\frac{x^{n+1}}{(1+\theta x)^{3/2}}.$$

Now, since $0 < \theta < 1$ and $-1 < x < 0$, we conclude that

$$0 < \frac{1-\theta}{1+\theta x} < 1 \quad\text{and}\quad \frac{1}{1+\theta x} < \frac{1}{1+x}.$$

Hence
$$|R_n(x)| < \frac{1\cdot 3\cdots(2n+1)}{n!2^{n+1}}\frac{|x|^{n+1}}{(1+x)^{3/2}}. \tag{9}$$

The argument from here on is much as it was in Example 1, but just a bit more involved. We let $c_n$ be the expression on the right in (9) and find that

$$c_{n+1} = \frac{2n+3}{2n+2}|x|c_n.$$

Now $|x| < 1$ and $\dfrac{2n+3}{2n+2}|x| \to |x|$. Hence if we choose some number $r$ such that $|x| < r < 1$, there is some value of $N$ so large that

$$\frac{2n+3}{2n+2}|x| < r \quad\text{if}\quad N \le n.$$

We then have $c_{n+1} < rc_n$ if $N \leq n$. Therefore $c_{N+k}$ decreases at least as fast as $r^k c_N$, and so must approach 0 as $k \to \infty$. But then $R_n(x) \to 0$.

If $m$ is a positive integer, the binomial theorem of algebra tells us that for any $x$,

$$(1 + x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \cdots + x^m. \tag{10}$$

The last term here has coefficient 1 as part of the general law of formation of the coefficients, because

$$\frac{m(m-1)\cdots[m-(m-1)]}{m!} = 1.$$

If $m$ is a number which is not 0 or a positive integer, the binomial theorem of algebra tells us nothing about $(1 + x)^m$. But we can apply Taylor's formula to this function, with $a = 0$ and $n$ taken to be any positive integer we please. It can be proved that $R_n(x) \to 0$ as $n \to \infty$ provided that $|x| < 1$. Examples 1 and 2 were illustrations of this. The proof works out with Lagrange's formula if $0 < x < 1$ and with Cauchy's formula if $-1 < x < 0$. To complete the discussion it is necessary to consider what happens for other values of $x$. It turns out that $R_n(x) \to 0$ if $x = 1$ and $m > -1$, and also if $x = -1$ and $m > 0$. We omit the proofs, which are a bit involved. The Taylor's series diverges if $|x| > 1$. Hence if $|x| < 1$ and in the indicated cases when $|x| = 1$, the Taylor's series formula is

$$(1 + x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \cdots + \frac{m(m-1)\cdots(m-n+1)}{n!}x^n + \cdots. \tag{11}$$

This turns out to be the same as the binomial formula (10) when $m$ is a positive integer. But otherwise the series (11) is genuinely an infinite series. It is called *the binomial series*.

**Example 3:** Use the binomial series to compute $\sqrt{2510}$ to four decimal places.

We write

$$\sqrt{2510} = \left[2500\left(1 + \frac{10}{2500}\right)\right]^{1/2} = 50\left(1 + \frac{4}{10^3}\right)^{1/2}.$$

Applying (11) with $x = 4/10^3$ and $m = \frac{1}{2}$, we obtain

$$50\left[1 + \frac{1}{2}\frac{4}{10^3} - \frac{1}{8}\frac{16}{10^6} + \frac{1}{16}\frac{64}{10^9} - \cdots\right] = 50 + \frac{1}{10} - \frac{1}{10^4} + \frac{2}{10^7} - \cdots$$

$$= 50.0999,$$

to four decimal places.

## EXERCISES

**1.** Write out the binomial series for each case. Simplify the coefficients as much as possible.

(a) $(1 + x)^{-2}$.
(b) $(1 - x)^{-3}$.
(c) $(1 + x)^{-3/2}$.
(d) $(1 + x)^{3/2}$.

**2.** Express each function in the form $C(1 + t)^m$, where $C$ is a constant and $t$ is some simple function of $x$. Then apply the binomial series and replace $t$ by its value in terms of $x$, thus getting a series formula for the given function. Indicate to what values $x$ must be restricted to make $|t| < 1$.

(a) $(25 + x^4)^{1/2}$.
(b) $(4 - x^2)^{-2}$.
(c) $(9 + x^2)^{-3/2}$.
(d) $(8 - x^3)^{1/3}$.

**3.** If $(A - A_0)/A = 3x$, where $|3x| < 1$, show that $A/A_0 = 1 + 3x + (3x)^2 + \cdots + (3x)^n + \cdots$.

**4.** The formula $D = y_2 - y_1$, where

$$y_2 = \sqrt{R^2 - x^2}, \qquad y_1 = h + \sqrt{(R - h)^2 - x^2}$$

arose in a situation where it was desirable to be able to compute $D$ readily from given values of $R$, $h$, and $x$. If $0 < h < R$ and if $x$ is small in comparison with $R$ and $R - h$, show that an approximate formula for $D$ is

$$D = \frac{hx^2}{2R(R - h)}.$$

The idea is to use the binomial series for the square roots, using the procedure suggested in Exercise 2. Take just two terms of each series.

**5.** (a) Write Taylor's formula and Lagrange's remainder for $f(x) = \sqrt{x}$, $a = 9$, $n = 2$.
(b) Estimate the size of $|R_2(x)|$ if $9 < x \leq 10$.
(c) Compute $\sqrt{10}$ to as many decimal places as is justified when $R_2(x)$ is neglected.

**6.** (a) Write Taylor's formula and Lagrange's remainder for $f(x) = x^{-1}$, $a = 10$, $n = 2$.
(b) Estimate the size of $|R_2(x)|$ if $10 < x \leq 11$.
(c) Compute $1/10.05$ to as many decimal places as is justified when $R_2(x)$ is neglected.

**7.** If $f(x) = (1 + x)^{1/2}$, and $a = 0$, show that $R_2(x)$ lies between

$$\frac{|x|^3}{16(1 + x)^{5/2}} \quad \text{and} \quad \frac{|x|^3}{16}$$

when $-1 < x$ and $x \neq 0$.

**8.** Give an estimate of $R_1(x)$ analogous to the statement in Exercise 7 for $f(x) = (1 + x)^{3/2}$, $a = 0$.

9. (a) Prove the Lagrange formula (7) for the case $n = 2$. Start with $R_2(x) = (x - a)^3M/3!$ and define a function $\phi$ by a formula analogous to (4). Then finish the proof. (b) Show how to give the proof of (7) for any $n$.

10. The law of the mean for a function $F$ can be written in the form

$$F(a + h) - F(a) = hF'(a + \theta h),$$

where $\theta$ is some number such that $0 < \theta < 1$. (a) To prove the Cauchy formula (8) when $n = 2$, let

$$F(u) = f(x) - f(u) - f'(u)(x - u) - f''(u)\frac{(x - u)^2}{2!}.$$

Note that $F(x) = 0$ and $F(a) = R_2(x)$. Show that

$$F'(u) = -\frac{f^{(3)}(u)}{2!}(x - u)^2.$$

Hence show that when we write $x - a = h$ and apply the law of the mean, the result is (8) with $n = 2$. (b) Generalize this argument for any $n$.

## 15-5 Absolute and Conditional Convergence

We now begin a systematic study of some important methods for determining whether a given series is convergent or divergent.

There is a convenient symbolism which enables us to abbreviate the writing of an infinite series. The series

$$a_1 + a_2 + \cdots + a_n + \cdots \tag{1}$$

is denoted by

$$\sum_{n=1}^{\infty} a_n, \quad \text{or sometimes by} \quad \sum a_n.$$

The Greek capital letter *sigma* that appears here is called a *summation sign.* There are occasions when we number the terms of a series 0, 1, 2, $\cdots$ instead of 1, 2, 3, $\cdots$. Also, we may drop off the first few terms of a given series. For instance, if we dropped off the first five terms of the series (1), our new series would be

$$\sum_{n=6}^{\infty} a_n.$$

In this case, of course, $a_n$ is no longer the $n$th term of the series.

The basic tests for convergence are built up from considering series with terms $a_n$ for all of which $a_n \geq 0$. However, we do often need to study series in which there are an infinite of negative terms as well as an infinite number of positive terms. Such series can sometimes be proved convergent by considering the series which is obtained from it by replacing each term by its absolute value. The basic fact here is stated as follows.

THEOREM 15-C. *If* $\sum |u_n|$ *is convergent, so is* $\sum u_n$.

*Proof.* The argument is based on Theorem 14-C (Cauchy's principle of convergence). Let us write

$$S_n = u_1 + \cdots + u_n, \qquad T_n = |u_1| + \cdots + |u_n|.$$

The hypothesis is that the sequence $\{T_n\}$ is convergent, and we wish to prove that $\{S_n\}$ is convergent. Now, if $m < n$,

$$S_n - S_m = u_{m+1} + u_{m+2} + \cdots + u_n,$$
$$T_n - T_m = |u_{m+1}| + |u_{m+2}| + \cdots + |u_n|.$$

But by the property of absolute values expressed in (2), § 14-1, we see that

$$|S_n - S_m| \leq |T_n - T_m|. \tag{2}$$

We now use Theorem 14-C. The convergence of $\{T_n\}$ implies that $|T_n - T_m| \to 0$ as $m$ and $n \to \infty$. By (2), then, $|S_n - S_m| \to 0$ also, and this implies that $\{S_n\}$ is convergent.

The converse of Theorem 15-C is false. It may well happen that $\sum u_n$ converges but $\sum |u_n|$ diverges. We have an example of this with the two series

$$1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} + \cdots, \qquad 1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \cdots.$$

The first of these we know to be convergent, with sum log 2 [see § 15-2, putting $x = -1$ in (4)]. The second is the harmonic series, which we know to be divergent (Exercise 6, § 15-1).

When a series $\sum u_n$ is such that $\sum |u_n|$ is convergent, the original series is said to be *absolutely convergent.* If a series is convergent, but not absolutely convergent, it is called *conditionally convergent.* In this case the fact that the series is convergent is due more to the effect of cancellation of terms of unlike sign than it is due to the diminution of the size of terms as $n \to \infty$.

## 15-6 Comparison Tests for Convergence

In this section we confine our attention to series with terms none of which are negative. Let $u_n$ be the $n$th term of such a series, and let $S_n$ be the sum of the first $n$ terms. Then $\{S_n\}$ is a nondecreasing sequence, as a result of the fact that $u_n \geq 0$ for every $n$. In this case then, we know from Chapter XIV that the series is convergent if the sequence $\{S_n\}$ has an upper bound, but not otherwise. Hence if we are to prove that the series is convergent, we must somehow demonstrate that the sequence has an upper bound.

If we have a stock of infinite series with positive terms, and for each of which we know whether it is convergent or divergent, this gives us a means of testing other series. The principle is this: Suppose we have two series $\sum a_n$ and $\sum b_n$, with $a_n \geq 0$, $b_n \geq 0$, and

$$A_n = a_1 + \cdots + a_n, \qquad B_n = b_1 + \cdots + b_n.$$

Suppose also that in some way we are able to show that $A_n \leq B_n$ for every $n$. Now each of the sequences $\{A_n\}$, $\{B_n\}$ is nondecreasing. If $\{B_n\}$ has an upper bound, then clearly so does $\{A_n\}$. On the other hand, if $\{A_n\}$ *does not* have an upper bound, neither does $\{B_n\}$. From these observations we conclude (always assuming that $A_n \leq B_n$):

$$\text{if } \sum b_n \text{ converges, so does } \sum a_n;$$

$$\text{if } \sum a_n \text{ diverges, so does } \sum b_n.$$

**Example 1:** We can use the known divergence of the harmonic series to prove that the series

$$\frac{1}{1^p} + \frac{1}{2^p} + \cdots + \frac{1}{n^p} + \cdots$$

is divergent if $p < 1$. Let $a_n = 1/n$, $b_n = 1/n^p$. Then $a_1 = b_1$ and $a_n < b_n$ if $n > 1$, so that $A_n \leq B_n$ for every $n$. The conclusion now follows by application of the foregoing remarks.

In trying to reduce the foregoing principle to a rule which is easily applicable, it is most practical to replace the comparison $A_n \leq B_n$ by a direct comparison of the terms of the two series. Evidently, if $a_n \leq b_n$ for every $n$, then $A_n \leq B_n$. Now, in practice, when we think of a series which it is natural to compare with another series, it often happens that we do not have exactly $a_n \leq b_n$ for every $n$. Instead it may turn out that $a_n \leq cb_n$ for every $n$, where $c$ is some positive constant. But this is just as useful for our purpose, for, if the series $\sum b_n$ is convergent, so is the series $\sum (cb_n)$, and if $a_n \leq cb_n$ for every $n$, then $\sum a_n$ is also convergent.

There is one more extension of the principle which is worth noting. Suppose that $a_n \leq cb_n$ is not true for *every* $n$, but that it *is* true for all $n$'s after a certain specific integer $N$. We can then make our comparison of the series which are obtained by discarding the first $N$ terms of each one. This dropping of the first $N$ terms has no effect on the matter of convergence or divergence.

We state all this as a formal theorem.

THEOREM 15-D. *Consider two series of nonnegative terms, with nth terms $a_n$ and $b_n$, respectively. Suppose there is some positive constant $c$ and some fixed $N$ such that $a_n \leq cb_n$ if $N \leq n$. Then, if $\sum b_n$ is convergent, so is $\sum a_n$, and if $\sum a_n$ is divergent, so is $\sum b_n$.*

It may be observed, as a matter of logic, that the second part of the proposition is just another way of stating the first part; i.e., if the truth of $P$ implies the truth of $Q$, then the falsity of $Q$ implies the falsity of $P$.

**Example 2:** The series

$$\frac{3}{1}\cdot 1 + \frac{5}{2}\frac{1}{2} + \frac{7}{3}\left(\frac{1}{2}\right)^2 + \cdots + \frac{2n+1}{n}\left(\frac{1}{2}\right)^{n-1} + \cdots$$

is convergent, because

$$\frac{2n+1}{n}\left(\frac{1}{2}\right)^{n-1} \leq 3\left(\frac{1}{2}\right)^{n-1}$$

and the geometric series with $n$th term $(\frac{1}{2})^{n-1}$ is convergent.

A person who works with infinite series a good deal learns to judge the probable behavior of series in a great many cases by looking at the "order of magnitude" of the $n$th term. That is, he tries to see if $a_n$ is approximately a constant multiple of something comparatively simple, such as

$$\frac{1}{n}, \quad \frac{1}{n^2}, \quad \frac{1}{n!}, \quad \text{or} \quad r^n.$$

He then expects the series to behave like the series with $n$th term of the simpler form. *This used only if the terms are all of one sign.* The mathematical justification of this approach to the examination of a series is contained in the next theorem.

THEOREM 15-E. *Suppose $\sum a_n$ and $\sum b_n$ are series whose terms are all positive. Suppose that*

$$\lim_{n\to\infty} \frac{a_n}{b_n} \text{ exists, } = C, \qquad \text{where } C \neq 0.$$

*Then, if one series converges, so does the other (and hence, if one series diverges, so does the other).*

For the proof, see Exercise 4.

As an illustration, consider the series in Example 2. Let

$$a_n = \frac{2n+1}{n}\left(\frac{1}{2}\right)^{n-1}, \qquad b_n = \left(\frac{1}{2}\right)^{n-1}.$$

Then $\lim_{n\to\infty} \frac{a_n}{b_n} = 2$. Since $\sum b_n$ converges, so does $\sum a_n$.

## EXERCISES

1. (a) Show that the series with $a_n = 1/n^n$ is convergent. (b) The same, if $a_n = 1/n!$.

2. (a) Show that the series with $a_n = 1/(2n - 1)$ is divergent. (b) The same, if $a_n = 1/2n$.

3. Determine the convergence or divergence of each series. Indicate your methods.

(a) $\Sigma \frac{n+1}{n!}$.

(b) $\Sigma \frac{n^2}{n^3+n+1}$.

(c) $\Sigma \frac{n^2}{(n+1)!}$.

(d) $\Sigma \frac{3n+1}{n} \frac{3^n}{4^n-1}$.

(e) $\sum_{n=2}^{\infty} \frac{1}{\log n}$.

(f) $\Sigma \frac{2^n+n}{3^n-n}$.

4. Prove Theorem 15-E. Begin by explaining why, under the given conditions, there is an $N$ such that

$$\frac{1}{2}C < \frac{a_n}{b_n} < \frac{3}{2}C$$

if $N \le n$. Then finish the argument, writing the reasoning out in full.

5. Prove that $\Sigma\,(1/n^2)$ is convergent by considering the partial sums $S_1$, $S_3$, $S_7$, $S_{15}$, $\cdots$, and showing by induction on $n$ that if $k = 2^n - 1$ then $S_k \le 1 + 1 + \frac{1}{2} + \cdots + 1/2^{n-1}$. Why does this imply convergence? Can you adapt this argument so as to prove $\Sigma\,(1/n^p)$ convergent when $p > 1$?

## 15-7 Improper Integrals and the Integral Test

Suppose $f$ is a function of $x$ which is defined and continuous when $x \ge 0$. The symbol

$$\int_0^\infty f(x)\,dx \tag{1}$$

is called an *improper* integral. To define the convergence or divergence of this integral we consider the function

$$F(t) = \int_0^t f(x)\,dx$$

and consider how $F(t)$ behaves as $t \to +\infty$. If $F(t)$ approaches a limit $I$ as $t \to +\infty$, we say that the integral in (1) is convergent and that its value is the number $I$. If $F(t)$ does not approach a limit, we call the integral in (1) divergent and do not speak about a value of it.

The lower limit of the integral need not be 0. It can be any number $a$, provided that $f$ is continuous when $x \ge a$.

**Example 1:** Consider $\int_1^\infty \frac{dx}{x^p}$, where $p \ne 1$. In this case

$$F(t) = \int_1^t \frac{dx}{x^p} = \frac{x^{-p+1}}{1-p}\bigg|_1^t = \frac{1}{1-p}(t^{-p+1} - 1).$$

When $t \to +\infty$, $F(t) \to 1/(p-1)$ if $p > 1$, but $F(t) \to +\infty$ if $p < 1$. Hence the integral is convergent, with value $1/(p-1)$ when $p > 1$. But it is divergent when $p < 1$. It is also divergent if $p = 1$, but logarithms are involved in that case. See Exercise 3.

**Example 2:** The integral $\int_0^\infty \sin x\, dx$ is divergent, because in this case $F(t) = 1 - \cos t$ approaches no limit as $t \to \infty$.

Our main need of improper integrals in this book is for their usefulness in connection with infinite series, as we shall see in a moment. However, there is another type of improper integral about which it is useful to know. It is defined, and a few things about it are considered, in Exercise 6.

*Estimating Sums by Integrals*

If $a_n$ depends on $n$ in a suitable way, it may be conveniently possible to estimate the value of the sum

$$S_n = a_1 + a_2 + \cdots + a_n \tag{2}$$

by means of an integral. Let us suppose we can find a positive continuous function $f$ of $x$ such that $f(n) = a_n$ for each $n$, and suppose that $f(x)$ decreases as $x$ increases. Then, for any positive integer $k$,

$$a_{k+1} = f(k + 1) < \int_k^{k+1} f(x)\, dx < f(k) = a_k. \tag{3}$$

If we write these inequalities down for $k = 1, 2, \cdots, n$ and add, we obtain

$$a_2 + \cdots + a_{n+1} < \int_1^{n+1} f(x)\, dx < a_1 + \cdots + a_n. \tag{4}$$

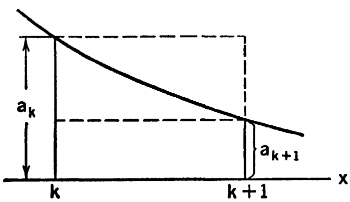


**Fig. 15-1**

The geometrical aspect of (3) is shown in Fig. 15-1. Evidently we could also get estimates in this way of the sum of any consecutive number of terms in the series, such as $a_{10} + \cdots + a_{25}$, say.

For the purpose of testing the convergence of the series, observe that (4) can be written as two inequalities:

$$S_{n+1} < a_1 + \int_1^{n+1} f(x)\, dx \tag{5}$$

and

$$\int_1^{n+1} f(x)\, dx < S_n. \tag{6}$$

From (5) we conclude that if the integral $\int_1^\infty f(x)\, dx$ is convergent, with a value $I$, then $a_1 + I$ is an upper bound for the sequence $\{S_n\}$, and hence the series $\Sigma\, a_n$ is convergent. On the other hand, if the integral diverges, this must be because

$$\int_1^t f(x)\, dx \to +\infty \quad \text{as} \quad t \to +\infty.$$

In that case (6) shows that the sequence $\{S_n\}$ has no upper bound, and the series diverges. We state these results formally.

THEOREM 15-F. *If $f$ is a positive continuous function of $x$ which decreases as $x$ increases, and if $f(n) = a_n$ for every $n$, then the series $\Sigma\, a_n$ and the integral $\int_1^\infty f(x)\,dx$ are either both convergent or both divergent.*

This theorem is the most convenient method of testing certain series.

**Example 3:** The series $\sum_2^\infty \frac{1}{n\,(\log n)^2}$ is convergent, for let us take $f(x) = \frac{1}{x\,(\log x)^2}$. In this case we must work when $x \geq 2$ instead of when $x \geq 1$, but that is an unimportant detail. The function $f$ has the requisite properties, and

$$\int_2^t \frac{dx}{x\,(\log x)^2} = -\frac{1}{\log x}\Big|_2^t = \frac{1}{\log 2} - \frac{1}{\log t}.$$

(We integrate by letting $u = \log x$.) It is now clear that the integral is convergent. Hence so is the series.

## EXERCISES

**1.** Examine each series for convergence or divergence, using the integral test.

(a) $\Sigma \frac{1}{n^2}$.

(b) $\Sigma \frac{1}{n^p}$.

(c) $\Sigma \frac{1}{2n - 1}$.

(d) $\sum_{n=2}^\infty \frac{1}{n \log n}$.

(e) $\Sigma \frac{1}{\sqrt{n^2 + 25}}$.

(f) $\Sigma \frac{n}{e^n}$.

**2.** Proceed as directed in Exercise 1.

(a) $\Sigma \frac{n}{n^4 + 1}$.

(b) $\Sigma \frac{1}{n(n + 1)}$.

(c) $\sum_{n=2}^\infty \frac{1}{n\,(\log n)^p}$.

(d) $\Sigma\, ne^{-n^2}$.

**3.** Discuss Example 1 in case $p = 1$.

**4.** If $S_n$ is the sum of the first $n$ terms of the harmonic series, show that $\log (n + 1) < S_n < 1 + \log n$.

**5.** Show that

$$\int_{100}^{1001} \frac{dx}{x^2} < \frac{1}{100^2} + \cdots + \frac{1}{1000^2} < \int_{99}^{1000} \frac{dx}{x^2}.$$

Work this out and obtain the upper and lower estimates as six-place decimals.

**6.** If $f(x)$ is continuous when $0 < x \leq a$, where $a > 0$, but $f(x)$ does not remain bounded as $x \to 0$, the integral of $f$ from 0 is $a$ is called improper.

It is called convergent if

$$\lim_{t \to 0} \int_t^a f(x)\, dx = I$$

exists, and in that case we write $I = \int_0^a f(x)\, dx$. If the indicated limit does not exist, the improper integral from 0 to $a$ is called divergent and we do not assign it a value. Similar definitions are made for the case in which $f$ is continuous when $a \le x < b$ but $f(x)$ does not remain bounded as $x \to b$.

Using these definitions, discuss the values of $p$ for which

$$\int_0^1 \frac{dx}{x^p} \quad \text{and} \quad \int_0^1 \frac{dx}{(1-x)^p} \quad \text{are convergent.}$$

## 15-8 Alternating Series

A series with terms $a_1, a_2, a_3, \cdots$ is called *alternating* if successive terms are always of opposite signs. For example, the series (10) expressing $\pi/4$ in § 15-2 is of this kind.

We consider just one very simple but also very important theorem about alternating series:

THEOREM 15-G. *Consider an alternating series with terms $a_n$ such that $|a_{n+1}| \le |a_n|$ for every $n$, and also such that $a_n \to 0$. Such a series is convergent.*

*Proof.* Let us consider the two kinds of partial sums: those with an odd number of terms, and those with an even number. We suppose, for definiteness, that the first term of the series is positive. Then, because the terms alternate in sign and never increase in absolute value, we see that $S_2 \le S_1$, $S_2 \le S_3$ but $S_3 \le S_1$, $S_4 \le S_3$ but $S_2 \le S_4$, and so on. The situation is shown in Fig. 15-2 on the assumption that $|a_{n+1}|$ is always definitely less than $|a_n|$. This shows up the fact that $S_2, S_4, S_6, \cdots$ form a bounded nondecreasing sequence which must converge to its least upper bound $S$, while $S_1, S_3, S_5, \cdots$ form a bounded nonincreasing sequence converging to its greatest lower bound. Since $S_{2n+1} - S_{2n} = a_{2n+1}$, a difference which approaches 0 as $n \to \infty$, it is clear that the greatest lower bound of the odd-index sequence is the same as the least upper bound of the even-index sequence. Hence this common value $S$ is the limit of the entire sequence $\{S_n\}$.

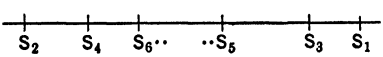


Fig. 15-2

It is also clear that the limit $S$ is bracketed by any two consecutive partial sums. This gives us:

THEOREM 15-H. *When the alternating series satisfies the conditions of Theorem* 15-G, *the sum $S$ of the series satisfies the inequality $|S - S_n| \le |a_{n+1}|$.*

This gives us an estimate of how close we are to $S$ when we use $S_n$ as an approximate value of $S$.

## EXERCISES

**1.** In each case determine whether or not Theorem 15-G is applicable to the series. The indicated expression is $|a_n|$ and it is assumed the signs alternate. In some cases a convenient test for $|a_{n+1}| \leq |a_n|$ is made by considering the ratio of consecutive terms. In other cases one may think of $n$ as a continuous variable and consider the derivative of $|a_n|$ with respect to $n$.

(a) $\dfrac{n+1}{2n}$.

(b) $\dfrac{1}{\log n}$.

(c) $\dfrac{2n+1}{n^2}$.

(d) $\dfrac{\sqrt{n}}{n+1}$.

(e) $\dfrac{n^2+n}{n^3+1}$.

(f) $\dfrac{\log n}{n}$.

(g) $ne^{-n}$.

(h) $\dfrac{\sqrt{n+1}}{n+\sqrt{n}}$.

**2.** Proceed as directed in Exercise 1.

(a) $\dfrac{1\cdot 3 \cdots (2n-1)}{2\cdot 4 \cdots (2n)} \dfrac{1}{n+1}$.

(b) $\dfrac{2\cdot 4 \cdots (2n)}{1\cdot 3 \cdots (2n-1)} \dfrac{1}{n^2}$.

(c) $\dfrac{n!}{1\cdot 3 \cdots (2n-1)}$.

(d) $\dfrac{1\cdot 3 \cdots (2n-1)}{(n+1)(n+2) \cdots (2n)}$.

**3.** Compute $e^{-0.1}$ by using three terms of the Taylor's series of $e^x$ with $a = 0$. Estimate the accuracy of the result.

**4.** Find the cosine of 1 radian, accurate to three decimal places, by using the Taylor's series of $\cos x$ with $a = 0$.

## 15-9 The Ratio Test

The ratio test, which is the subject of this section, is founded on two things: (1) a geometric series with $n$th term $cr^n$ is convergent if $|r| < 1$; (2) the result stated in Theorem 15-D. We also appeal to Theorem 15-C. The ratio test is a test which can be applied to the absolute values of the terms of some given series.

THEOREM 15-I. *From the infinite series* $\Sigma\, u_n$, *where each* $u_n \neq 0$, *form the ratio* $u_{n+1}/u_n$ *and find the limit of the absolute value of this ratio (we assume the limit exists):*

$$\lim_{n\to\infty} \left|\frac{u_{n+1}}{u_n}\right| = t. \tag{1}$$

*Then the given series converges absolutely if $t < 1$. If $t > 1$, the given series is divergent. If $t = 1$, the test gives no information.*

*Proof.* The idea of the proof is that when $t < 1$ we can compare $|u_n|$ with the $n$th term of a convergent geometric series. The precise argument runs as follows. Choose a number $r$ between $t$ and 1. Then since the ratio $|u_{n+1}|/|u_n|$ approaches $t$, it must become and remain less than $r$ when $n$ is sufficiently large, say when $N \leq n$. We then have

$$|u_{N+1}| \leq r|u_N|, \qquad |u_{N+2}| \leq r|u_{N+1}| \leq r^2|u_N|,$$

and by induction, $|u_{N+k}| \leq r^k|u_N|$. Now we can make our comparison with a convergent geometric series, and we conclude that the original series converges absolutely.

When $t > 1$, the limit (1) shows that we have $|u_{n+1}| > |u_n|$ when $n$ is sufficiently large. This prevents the terms from approaching 0, and hence the series diverges, by Theorem 15-A.

No conclusions can be drawn if $t = 1$. Examples with $t = 1$ may be given to show that the series may be absolutely convergent, but may also be conditionally convergent, or even divergent. See Exercise 1.

It may happen in particular cases that $t = 1$ but that also

$$\left|\frac{u_{n+1}}{u_n}\right| \geq 1$$

for all values of $n$. In this case the original series must be divergent, because under the conditions here stated the terms cannot approach 0. On the other hand, if $t = 1$ and $|u_{n+1}| < |u_n|$ for every $n$, it is *not* possible to conclude from this that the series converges (consider the harmonic series, for instance).

The ratio test is most often applied to series in which $u_n$ is of the form $a_nx^n$. Such a series is called a *power series* in $x$. In a typical application to such a series, we conclude that the series converges absolutely for certain values of $x$, diverges for other values of $x$, and we are left with two values of $x$ (always of the type $x = \pm c$) for which the ratio test is indecisive because $t = 1$. For these values of $x$ the convergence or divergence of the series must be decided by some other methods. Sometimes we can use the alternating series test of Theorem 15-G.

**Example:** Consider the series

$$x - \frac{x^2}{2\cdot 5} + \frac{x^3}{3\cdot 5^2} - \cdots + (-1)^{n+1}\frac{x^n}{n\cdot 5^{n-1}} + \cdots. \tag{2}$$

We begin the study of this problem by using the ratio test. For this purpose we assume that $x \neq 0$, since the series plainly converges and does not need to be tested when $x = 0$. We set

$$u_n = (-1)^{n+1}\frac{x^n}{n\cdot 5^{n-1}}, \qquad u_{n+1} = (-1)^{n+2}\frac{x^{n+1}}{(n+1)5^n},$$

$$\frac{u_{n+1}}{u_n} = \frac{(-1)^{n+2}x^{n+1}}{(n+1)5^n}\,\frac{n5^{n-1}}{(-1)^{n+1}x^n} = -\frac{n}{n+1}\frac{x}{5}.$$

Then
$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \lim_{n\to\infty}\frac{n}{n+1}\frac{|x|}{5} = \frac{|x|}{5}.$$

In this case the $t$ of the ratio test is $|x|/5$. Therefore we conclude that the series is absolutely convergent if $|x| < 5$ (i.e., $-5 < x < 5$), and divergent if $|x| > 5$ (i.e., $x < -5$ or $5 < x$). If $|x| = 5$ (i.e., if $x = 5$ or $-5$) we cannot conclude anything about convergence or divergence of the series by the ratio test. Therefore, when $x = \pm 5$ we must test the series in some other way. The first step is to put the values $x = 5, -5$ into the series and see what we get. When $x = 5$ the series becomes

$$5 - \frac{5}{2} + \frac{5}{3} - \cdots + (-1)^{n+1}\frac{5}{n} + \cdots.$$

This is an alternating series that satisfies the conditions of Theorem 15-G, and is therefore convergent. If $x = -5$, series (2) becomes

$$-5 - \frac{5}{2} - \frac{5}{3} - \cdots - \frac{5}{n} - \cdots.$$

This series is divergent, because it is obtained by multiplying each term of the harmonic series by $-5$.

Our solution is now complete. The series (2) converges if $-5 < x \le 5$, and diverges for all other values of $x$.

## EXERCISES

**1.** (a) Apply Theorem 15-I to the series $\sum_{n=1}^{\infty} \frac{x^n}{n^p}$ and state the conclusions. Note that they do not depend on $p$. (b) Discuss the convergence or divergence in the cases where the ratio test fails. Make a complete catalogue of conclusions, depending on the two values of $x$ and various values of $p$.

**2.** Apply the ratio test in each case and state what you can conclude without applying any other test.

(a) $\sum \frac{n(n+1)3^{n+1}}{4^n}$.

(b) $\sum (-1)^n \frac{10^{2n}}{n!}$.

(c) $\sum (-1)^n \frac{n!}{2^n}$.

(d) $\sum (-1)^n \frac{n^2+1}{n^3}$.

(e) $\sum (-1)^n \frac{n^3 2^n}{3^n}$.

(f) $\sum n!x^n$.

(g) $\sum \frac{2^{2n}}{2\cdot 4 \cdots (2n)}$

(h) $\sum \frac{n!}{n^n}$.

**3.** In each case find all the values of $x$ for which the series is convergent.

(a) $\Sigma \dfrac{nx^n}{3^n}$.

(b) $\Sigma (-1)^{n+1} \dfrac{x^{2n}}{(2n)!}$.

(c) $\Sigma (-1)^n \dfrac{x^n}{n}$.

(d) $\Sigma n^2 \dfrac{(x+1)^n}{4^n}$.

(e) $\Sigma (-1)^{n+1} \dfrac{x^n}{n2^n}$.

(f) $\Sigma \dfrac{x^n}{\sqrt{n}}$.

(g) $\Sigma \dfrac{x^n}{2\cdot 4 \cdots (2n)}$.

(h) $\Sigma \dfrac{3^n}{n!x^n}$.

**4.** Proceed as directed in Exercise 3.

(a) $\Sigma \dfrac{x^n}{n(n+1)}$.

(b) $\Sigma (-1)^{n+1} \dfrac{(x-3)^n}{(2n+1)5^n}$.

(c) $\Sigma (-1)^n \dfrac{x^{2n+1}}{(2n+1)!}$.

(d) $\Sigma \dfrac{n}{2^n x^n}$.

(e) $\Sigma (-1)^n \dfrac{1\cdot 3 \cdots (2n-1)(3x)^{2n+1}}{2\cdot 4 \cdots (2n)(2n+1)}$.

(f) $\Sigma \dfrac{1}{n}\left(\dfrac{x-1}{x}\right)^n$.

(g) $\Sigma \dfrac{(x+2)^n}{n\cdot 2^{n-1}}$.

(h) $\sum_{n=2}^{\infty} \dfrac{\log n}{n} x^n$.

**5.** Suppose $R_n$ is the remainder after the term $u_n$ in a series, and that $\left|\dfrac{u_{n+1}}{u_n}\right| \le r < 1$. Show that $|R_n| \le \dfrac{r|u_n|}{1-r}$. In particular, if $r \le \frac{1}{11}$, this means that $|R_n| \le \dfrac{1}{10}|u_n|$. Use this result to compute the sum of the series

$$\sum_{n=1}^{\infty} \frac{1}{n\cdot n!2^n}$$

with an error less than 0.0005. How many terms are required?

**6.** The series for $(1+x)^{-1/2}$ is

$$1 - \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 - \cdots + (-1)^n \frac{1\cdot 3 \cdots (2n-1)}{2\cdot 4 \cdots (2n)} x^n + \cdots.$$

In order to discuss the convergence of this series completely it is necessary to have some motion of the size of the coefficients

$$a_n = \frac{1\cdot 3 \cdots (2n-1)}{2\cdot 4 \cdots (2n)}.$$

It is easy to see that $a_n \ge 1/2n$, but less easy to see that $a_n < (2n+1)^{-1/2}$. However, this is the same as

$$a_n < \frac{1}{a_n}\frac{1}{2n+1},$$

and in this form it is not so difficult to verify. Now discuss the given series.

## 15-10 Power Series

A series of the form

$$c_0 + c_1(x - a) + \cdots + c_n(x - a)^n + \cdots \tag{1}$$

is called a power series in $(x - a)$. The Taylor's series for a function is a power series. The constants $c_0, c_1, \cdots$ are called the coefficients of the power series. The nature of a power series is such that one of three things happens:

(a) The power series (1) may converge for all values of $x$.

(b) It may diverge except when $x = a$.

(c) There is a number $r > 0$ such that the series converges absolutely if $|x| < r$ and diverges if $|x| > r$. In this case a variety of behaviors are possible when $x = r$ or $x = -r$.

We shall not prove this assertion about the three-way alternative, but we observe that the exercises of § 15-9 illustrate the alternatives.

When the power series (1) converges, let us denote its value by $f(x)$. In this way the power series defines a function. Sometimes we start with a function and show that for certain values of $x$ it can be expressed as a power series. This was illustrated in § 15-2 and in the discussion of Taylor's series. But sometimes we define functions directly by constructing a power series and showing that it converges for certain values of $x$. This is done a great deal in more advanced mathematics.

In working with functions defined by power series there are several important things which it is useful to know. We state these things without proof. The proofs are ordinarily considered in a more advanced course in calculus.

If a function $f(x)$ is equal to the power series (1) in some interval centered at $x = a$, then this is the only such power series formula for $f(x)$ in this interval. That is, the coefficients are uniquely determined. The practical advantage of this is that, although the coefficients $c_n$ are given, as in Taylor's series, by the formulas

$$c_0 = f(a), \qquad c_n = \frac{f^{(n)}(a)}{n!}, \tag{2}$$

we do not always have to compute them in this way.

If the series (1) for $f(x)$ converges when $|x - a| < r$, then we can differentiate the series term-by-term to get derivatives of $f(x)$:

$$f'(x) = c_1 + 2c_2(x - a) + 3c_3(x - a)^2 + \cdots,$$

$$f''(x) = 2c_2 + 3 \cdot 2c_3(x - a) + \cdots,$$

and so on. These new series will also be convergent when $|x - a| < r$. We can also integrate from $a$ to $x$ if $|x - a| < r$.

$$\int_a^x f(t)\,dt = c_0(x - a) + \frac{1}{2}c_1(x - a)^2 + \frac{1}{3}c_2(x - a)^3 + \cdots.$$

Integration of a power series has two obvious direct applications. If we already have the power series for $f(x)$, it is sometimes far easier to integrate this series than it is to compute the series for the integrated function by using the formulas for the coefficients in Taylor's series. This is illustrated by the derivation of the series for $\tan^{-1} x$ in powers of $x$ from the series

$$\frac{1}{1 + x^2} = 1 - x^2 + x^4 - \cdots$$

which we get directly as a geometric series. For other examples see Exercises 1, 5. Another application of integration of power series is in the calculation of integrals that cannot be worked out by elementary anti derivatives.

**Example 1:** Compute the integral

$$\int_0^{1.5} \frac{dx}{\sqrt{16 + x^4}}.$$

First we must get a power series for the integrand. We do this by using the binomial series for $(1 + t)^{-1/2}$ (see Exercise 6, § 15-9):

$$(16 + x^4)^{-1/2} = \frac{1}{4}\left(1 + \frac{x^4}{16}\right)^{-1/2} = \frac{1}{4}\left[1 + \left(\frac{x}{2}\right)^4\right]^{-1/2}$$
$$= \frac{1}{4}\left[1 - \frac{1}{2}\left(\frac{x}{2}\right)^4 + \frac{1\cdot 3}{2\cdot 4}\left(\frac{x}{2}\right)^8 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\left(\frac{x}{2}\right)^{12} + \cdots\right].$$

This is valid if $|x| \leq 2$. In our integration $x$ goes from 0 to 1.5, so we may integrate the series. The result is

$$\int_0^{1.5} \frac{dx}{\sqrt{16 + x^4}} = \frac{1}{4}\left[x - \frac{1}{5}\left(\frac{x}{2}\right)^5 + \frac{1\cdot 3}{2\cdot 4}\cdot\frac{2}{9}\left(\frac{x}{2}\right)^9 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\frac{2}{13}\left(\frac{x}{2}\right)^{13} + \cdots\right]_0^{3/2}$$
$$= \frac{1}{4}\left[\frac{3}{2} - \frac{1}{5}\left(\frac{3}{4}\right)^5 + \frac{1}{12}\left(\frac{3}{4}\right)^9 - \frac{5}{104}\left(\frac{3}{4}\right)^{13} + \cdots\right].$$

The final computation is most expeditiously done with logarithms. The final value found in this way is approximately 0.365.

There are many ways of combining known series to get new series. We illustrate three different ways. Proofs of the legitimacy of these procedures are beyond the scope of this book.

Two power series may be multiplied together to obtain a new power series. The multiplication process is carried out by forming all possible products of terms of one series by those of the other and arranging the results according to powers of $x$. The resulting series is the power series for the product of the functions represented by the two original series.

**Example 2:** Find the Maclaurin series for $e^x \sin x$. We know that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots.$$

Therefore,

$$e^x \sin x = \left(1 + x + \frac{x^2}{2!} + \cdots\right)\left(x - \frac{x^3}{3!} + \cdots\right).$$

The work is arranged as follows:

$$\begin{array}{llllll} e^x \sin x = x & & -\frac{x^3}{3!} & & +\frac{x^5}{5!} & - \cdots \\ & +x^2 & & -\frac{x^4}{3!} & & +\frac{x^6}{5!} \; \cdots \\ & & +\frac{x^3}{2!} & & -\frac{x^5}{2!3!} & + \cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \end{array}$$

$$e^x \sin x = x + x^2 + \tfrac{1}{3}x^3 + \cdots.$$

It may not be easy to recognize the formula for the general term in this kind of process.

Two power series may be divided by the procedure used in dividing polynomials in algebra. The terms are arranged according to ascending powers. The result will be a series for the quotient. In particular cases either or both of the given series may be a polynomial.

**Example 3:** Find the power series in $x$ for

$$\frac{x^2}{1 - x + x^2 - x^3}.$$

The division process appears as follows:

$$\begin{array}{r|l} & x^2 + x^3 + x^6 + x^7 + \cdots \\ \hline 1 - x + x^2 - x^3 & x^2 \\ & \underline{x^2 - x^3 + x^4 - x^5} \\ & \quad x^3 - x^4 + x^5 \\ & \quad \underline{x^3 - x^4 + x^5 - x^6} \\ & \qquad\qquad x^6 \\ & \qquad\qquad \underline{x^6 - x^7 + x^8 - x^9} \\ & \qquad\qquad\quad x^7 - x^8 + x^9 \\ & \qquad\qquad\quad \cdots\cdots\cdots\cdot \end{array}$$

The result is

$$\frac{x^2}{1 - x + x^2 - x^3} = x^2 + x^3 + x^6 + x^7 + \cdots.$$

Still another useful procedure for finding power series representations is that of *substitution.*

**Example 4:** Find the terms through $x^3$ in the Maclaurin series for $\log(1 + \sin x)$.

We start from the series

$$\log(1+h) = h - \frac{1}{2}h^2 + \frac{1}{3}h^3 - \cdots,$$

and put

$$h = \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots.$$

Thus

$$\log(1+\sin x) = \left(x - \frac{x^3}{3!} + \cdots\right) - \frac{1}{2}\left(x - \frac{x^3}{3!} + \cdots\right)^2 + \frac{1}{3}\left(x - \frac{x^3}{3!} + \cdots\right)^3 - \cdots.$$

We now start the squaring and cubing of the series as indicated, and rearrange the terms according to ascending powers of $x$. The result, including all terms of degree three or less, is

$$\log(1+\sin x) = x - \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots.$$

The use of series is often convenient as an alternative to the use of l'Hospital's rule.

**Example 5:** Find $\lim_{x \to 0} \frac{x \cos x - \sin x}{x^2 \tan x}$.

The first step is to express the numerator and denominator in terms of infinite series.

$$\begin{aligned} x\cos x - \sin x &= x\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots\right) - \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right) \\ &= -\frac{x^3}{3} + \frac{x^5}{30} - \cdots; \end{aligned}$$

$$x^2 \tan x = x^2\left(x + \frac{1}{3}x^3 + \cdots\right) = x^3 + \frac{x^5}{3} + \cdots.$$

Then

$$\frac{x\cos x - \sin x}{x^2 \tan x} = \frac{-\frac{x^3}{3} + \frac{x^5}{30} - \cdots}{x^3 + \frac{x^5}{3} + \cdots} = \frac{-\frac{1}{3} + \frac{x^2}{30} - \cdots}{1 + \frac{x^2}{3} + \cdots}.$$

At the last step we cancelled a common factor $x^3$ from the numerator and denominator. Now we let $x \to 0$, and get

$$\lim_{x \to 0} \frac{x\cos x - \sin x}{x^2 \tan x} = -\frac{1}{3}.$$

This use of power series depends on the fact that a function defined by a power series in $x$ is continuous at $x = 0$.

## EXERCISES

**1.** Write a power series formula for $(1 - t^2)^{-1/2}$ and integrate it to get the Maclaurin series for $\sin^{-1} x$ (i.e., the Taylor's series in powers of $x$).

**2.** Differentiate the geometric series for $(1 - x)^{-1}$, and so obtain a series for $(1 - x)^{-2}$. Verify that it is the same as the binomial series for $(1 + h)^m$ with $h = -x$ and $m = -2$.

**3.** Differentiate the Maclaurin series for $\sin x$ and compare the result with the Maclaurin series for $\cos x$.

**4.** Verify that differentiation of the binomial series for $2(1 + x)^{1/2}$ gives the binomial series for $(1 + x)^{-1/2}$.

**5.** Obtain the series

$$\log (x + \sqrt{1 + x^2}) = \sinh^{-1} x$$
$$= x - \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7} + \cdots$$

by integration of another series.

**6.** Find a power series in $x$ for

(a) $\displaystyle\int_0^x \frac{\sin t}{t}\,dt.$

(b) $\displaystyle\int_0^x e^{-t^2}\,dt.$

(c) $\displaystyle\int_0^x \sqrt{1 + t^3}\,dt.$

**7.** Integrate the Maclaurin series for $\tan x$ (see Exercise 11, § 15-3) and obtain a series for $\log \cos x$.

**8.** Carry the series for $e^x \sin x$ through the term in $x^6$ (see Example 2).

**9.** Obtain the series

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

by long division.

**10.** Show that

$$\cos^2 x = 1 - x^2 + \frac{1}{3}x^4 - \cdots.$$

**11.** Show that $\dfrac{1 - x}{1 - x + x^2} = 1 - x^2 - x^3 + x^5 + x^6 - x^8 - x^9 + \cdots.$

**12.** Find the series

$$\operatorname{sech} x = \frac{2}{e^x + e^{-x}} = 1 - \frac{x^2}{2} + \frac{5}{24}x^4 - \cdots$$

by long division. What is the next nonvanishing term?

**13.** Calculate the Maclaurin series for $\tan x$ from the series for $\sin x$ and $\cos x$.

**14.** Find the Maclaurin series of each of the following functions, using either multiplication and binomial series, or long division.

(a) $\dfrac{x}{1-x}\dfrac{x}{1+x^2}$.

(b) $\dfrac{3-x}{2-x}\dfrac{1}{(1-x)^2}$.

(c) $\dfrac{1-x}{(2+x)^2}$.

(d) $\dfrac{(1-4x)^2}{1-2x}$.

**15.** Find the Maclaurin series for $(1+x+x^2)^{-2}$ by putting $h = x + x^2$ in the binomial series for $(1+h)^{-2}$.

**16.** Find the Maclaurin series for $(1-2ax+x^2)^{-1/2}$ through the term in $x^2$.

**17.** Carry the series in Example 4 through the term in $x^5$.

**18.** Put $h = 1 - \cos x$ in the series for $\log(1+h)$, and so find the Maclaurin series of $\log(2-\cos x)$.

**19.** Verify the following approximate values.

(a) $\displaystyle\int_0^1 (8+x^3)^{1/3}\,dx = 2.020.$

(b) $\displaystyle\int_0^{1/2} \frac{e^x-1}{x}\,dx = 0.570.$

(c) $\displaystyle\int_0^1 \cos\sqrt{x}\,dx = 0.764.$

**20.** Find the limits of the following expressions as $x \to 0$.

(a) $\dfrac{e^x - \cos x}{\sin x}$.

(b) $\dfrac{\sin x - \tan x}{\sin^2 x}$.

(c) $\dfrac{(\sin x + \tan x)^2}{\cosh x - 1}$.

(d) $\dfrac{\sqrt{1+x^2} + \cos x - 2}{x^4}$.

## Review Questions and Problems for Chapters XIII, XIV, XV

### CONCEPTS AND DEFINITIONS

**1.** Explain two aspects of the concept of a vector in the $xy$-plane.

**2.** Define two algebraic processes involving vectors and explain their geometrical aspect. How are vectors represented in terms of the two standard unit vectors?

**3.** Define the velocity and acceleration vectors, (a) in terms of the algebraic representation of the position vector, using the standard unit vectors, and (b) directly in terms of derivatives of a vector function.

**4.** State Newton's law in vector form.

**5.** Define the unit vectors **T** and **N**, and the tangential and normal components of acceleration.

**6.** Define curvature, radius of curvature, center of curvature, and the evolute of a curve.

**7.** Give the basic rules for working with inequalities.

**8.** What can be said about the absolute values of $ab$ and $a + b$?

9. Define what is meant by a section in the real number system.

10. What is meant by saying that the real number system is complete?

11. What does it mean to say that a sequence is bounded?

12. What is a monotonic sequence? What special thing of great importance can be said about bounded monotonic sequences?

13. Define the meaning of the statement $\lim_{n \to \infty} x_n = c$.

14. How is the limit concept for sequences used to define convergence of an infinite series?

15. Write down Taylor's formula with remainder, without specifying any particular way of expressing the remainder. Under what conditions on the remainder is the function represented by Taylor's *series?* How many particular formulas for $R_n(x)$ do you know?

16. Explain the terms absolute and conditional convergence.

17. What is the basic comparison principle which is used in testing for convergence or divergence of a series? What relation does it bear to upper bounds?

18. Explain the meaning of $\int_0^\infty f(x)\,dx$.

19. What is an alternating series? Are all such series convergent? Justify your answer.

20. What is a power series?

THEORY

1. If $\mathbf{R}$ is a vector function of $t$, explain how $d\mathbf{R}/dt$ is defined. Deduce the formula for $d(u\mathbf{R})/dt$ if $u$ and $\mathbf{R}$ are differentiable functions of $t$.

2. If $C$ is a curve, $\mathbf{R}$ is the position vector to a point on it, and $s$ is arc length along $C$, explain why $d\mathbf{R}/ds$ is a unit vector which is tangent to the curve if based at the tip of $\mathbf{R}$.

3. Give two different explanations (one analytical, one at least partly geometrical) of the fact that the length of the velocity vector is $|\mathbf{V}| = |ds/dt|$.

4. For uniform motion in a circular path with center at $O$, prove that the acceleration vector is opposite in direction to the position vector. What is the magnitude of the acceleration?

5. Prove the formula $A_T = d^2s/dt^2$ by two methods, and from one of these methods obtain also a formula for $A_N$.

6. Deduce the formula for $K$ for curves expressed in the form $y = f(x)$; for curves expressed in parametric form.

7. Work out the formulas for radial and transverse components of velocity and acceleration, starting from $\mathbf{R} = r\mathbf{u}_r$.

8. Explain why, if a mass particle at $P$ moves under the gravitational attraction of a fixed mass at $O$, the radius $OP$ sweeps out area at a constant rate.

9. Work out the inequalities relating the sizes of $|uv - u_0v_0|$ and $\left|\frac{u}{v} - \frac{u_0}{v_0}\right|$ to the sizes of $|u - u_0|$, $|v - v_0|$. For what limit theorems are these inequalities used?

10. Prove by inequalities that $\lim_{n\to\infty} \frac{4}{2n^2 - 1} = 0$ by showing how to find the $N$ for a given $\epsilon$.

11. State Cauchy's convergence principle and explain how to prove it by constructing a section in the real number system.

12. Try to devise an example of a convergent sequence, not monotonic, for which it is not easy to know what the limit is, but which can be asserted to be convergent as a consequence of the Cauchy convergence principle.

13. State a simple and important condition which is necessary but not sufficient for the convergence of an infinite series. Prove the necessity of the condition.

14. If Taylor's formula with remainder is applied to a polynomial of degree $k$, what happens to the remainder when $n \geq k$?

15. Show how to derive at least one of the forms of $R_n(x)$ in Taylor's formula.

16. Explain precisely a meaning for this statement: "$a_n$ is of the same order of magnitude as $b_n$ when $n$ is large, and therefore the two series $\Sigma\, a_n$ and $\Sigma\, b_n$ have the same behavior as regards convergence or divergence." Does this apply to *all* pairs of series?

17. Explain the use of improper integrals in the consideration of convergence of infinite series. Work out the basic inequalities on which the reasoning turns.

18. State the ratio test and explain the basic principle of the test insofar as it is used to prove convergence.

19. Assuming the legitimacy of certain procedures as described in the text, prove that if $f(x)$ is defined as the sum of a convergent series of powers of $x - a$ (say when $|x - a| < r$), then the power series is the Taylor's series of the function.

## PROBLEMS

1. Show that, if $\mathbf{R} = f(t)\mathbf{i} + g(t)\mathbf{j}$ has constant length and $f$ and $g$ are differentiable, then $\mathbf{R}$ and $d\mathbf{R}/dt$ are in general perpendicular.

2. Three points $P_1$, $P_2$, $P_3$ are given, forming a triangle. Let $Q_1$ be the midpoint of the side opposite $P_1$, with $Q_2$ and $Q_3$ defined similarly. Let $\mathbf{A}_1$ be the vector with direction and magnitude of the directed line $P_1Q_1$, and define $\mathbf{A}_2$ and $\mathbf{A}_3$ likewise. Show that $\mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3 = 0$.

**3.** Locate the positive value of $x$ for which the radius of curvature of the curve $b^2y = 3a^2x - x^3$ is least. Sketch the curve. Does the value of $x$ in question correspond to the relative maximum point on the curve?

**4.** Let $x_n = \left(1 + \frac{1}{n}\right)^n$, $n = 1, 2, \cdots$. Re-read § 8-4, including the second exercise, and then answer: What is the limit of the sequence $\{x_n\}$ and in what manner does $x_n$ approach the limit?

The information developed in the answer to Problem 4 plays a role in Problems 5, 6 and 7.

**5.** If $x_n = n^n/n!$, is the sequence $\{x_n\}$ monotonic? Is it bounded?

**6.** Prove that $\{x_n\}$ is convergent if $x_n = n^n/n!e^n$. It can be shown that the limit is 0.

**7.** If $x_n = (\log n)/n$, show that $\{x_n\}$ is ultimately monotonic by showing that $x_n > x_{n+1}$ is equivalent to $n > \left(1 + \frac{1}{n}\right)^n$. What is the limit of the sequence?

**8.** Let $\{x_n\}$ be a sequence defined as follows: $x_n = a_1 + a_2 + \cdots + a_n$, where the following information about the $a_n$'s is given. For each $n$, $a_n$ is a certain one of the four possible things,

$$\frac{1}{2^n}, \quad -\frac{1}{3^n}, \quad \left(\frac{2}{3}\right)^n, \quad -\left(\frac{3}{4}\right)^n.$$

Thus, for example, $a_1$ might be $-\frac{1}{3}$, $a_2$ might be $(\frac{2}{3})^2$, and so on. The rule is definite, but it is not known to you. Can you nevertheless prove that the sequence $\{x_n\}$ is convergent? If $x$ is the limit, can you estimate how far, *at most*, $x_m$ is from $x$?

**9.** When a uniform circular metal plate of radius $R$ is clamped around the edges and subjected to a normal force $P$ at its center, the deflection $w$ at any point whose distance from the center is $r$ is given by

$$w = \frac{P}{16\pi D}\left[2r^2 \log \frac{r}{R} + R^2 - r^2\right],$$

where $D$ is the flexural rigidity of the plate (a constant). Plot $w$ as a function of $r$, $0 < r \leq R$, and find the limiting values of $w$ and $dw/dr$ as $r \to 0$. Find the ratio $r/R$ at the point of inflection.

**10.** In a circle of variable radius $r$ a chord $BC$ subtends a variable central angle $\theta$ in such a way that the shorter arc $BC$ has a constant length $L$. Let $A$ be the smaller of the two areas into which the chord divides the circle. Express the area $A$ as a function of $\theta$ as the only variable. What happens to $A$ as $r \to \infty$ and $\theta \to 0$? What happens to $dA/d\theta$? Draw a graph of $A$ as a function of $\theta$, $0 < \theta \leq \pi$.

**11.** Show that the approximate formula $(1 + x)^{3/2} = 1 + \frac{3}{2}x$ is accurate to at least three decimal places if $-0.03 \leq x \leq 0$.

**12.** Show that the approximate formula $(1 + x)^{-1/2} = 1 - \frac{1}{2}x$ is accurate to at least two decimal places if $-0.1 \leq x \leq 0$.

**13.** In the formula $(1 + x)^{100} = 1 + 100x + 4950x^2 + R_2(x)$, find Lagrange's formula for $R_2$ and estimate the size of $|R_2|$ if $-0.001 \leq x < 0$. Find $(0.999)^{100}$ to the accuracy that is justified by neglect of $R_2$.

**14.** Suppose that $(1 + x)^{50}$ is computed approximately by using the binomial expansion and neglecting the terms involving powers of $x$ greater than the third. Use Lagrange's form of the remainder to obtain an expression for the error thus committed. (a) If $-0.01 \leq x \leq 0$, show that the error is less than 0.003. Hence compute $(0.99)^{50}$ to two places of decimals. (b) If $0 \leq x \leq 0.01$, show that the *percentage* error does not exceed $\frac{1}{4}\%$.

**15.** In the Maclaurin expansion of $\log(1 + x)$ show that

$$\frac{x^2}{2(1+x)^2} < R_1 < \frac{x^2}{2} \quad \text{if} \quad 0 < x.$$

**16.** (a) Write Taylor's formula with $a = \pi/3$, $n = 3$, for $\sin x$ and $\cos x$.
(b) Compute $\sin 61°$ by the result in (a), neglecting $R_3$. Show that $R_3$ is too small to affect the fourth decimal place. Use the approximation $\pi/180 = 0.0175$.

**17.** If $\dfrac{1}{d} = \dfrac{1}{p} + \dfrac{1}{p+h}$, where $h$ is small in comparison with $p$, expand $d$ in powers of $h/p$, and thus show that, approximately $d = \dfrac{1}{2}\left(p + \dfrac{h}{2}\right)$.

**18.** Show that, when $b$ is large and $h$ is small, $\dfrac{1}{b^2} - \dfrac{1}{(b+h)^2} = \dfrac{2h}{b^3}$, if we neglect higher powers of $h$.

**19.** The integrals

$$K = \int_0^{\pi/2} (1 - k^2 \sin^2 t)^{-1/2}\, dt, \qquad E = \int_0^{\pi/2} (1 - k^2 \sin^2 t)^{1/2}\, dt,$$

in which $0 < k < 1$, are known as the complete elliptic integrals of first and second kind, respectively. By using the binomial series for $(1 - x)^{-1/2}$ and $(1 - x)^{1/2}$, respectively, putting $x = k^2 \sin^2 t$ and integrating, show that

$$K = \frac{\pi}{2}(1 + a_1^2 k^2 + \cdots + a_n^2 k^{2n} + \cdots),$$

$$E = \frac{\pi}{2}\left(1 - a_1^2 \frac{k^2}{1} - a_2^2 \frac{k^4}{3} - \cdots - a_n^2 \frac{k^{2n}}{2n-1} - \cdots\right),$$

where $a_n = \dfrac{1\cdot 3 \cdots (2n-1)}{2\cdot 4 \cdots 2n}$.

**20.** Show that the total perimeter of an ellipse of major axis $2a$ and eccentricity $k$ is $4aE$ (see the previous problem). If powers of $k$ higher than 2 are neglected, what is the total perimeter of the ellipse $b^2x^2 + a^2y^2 = a^2b^2$, where $a > b$?

*CHAPTER XVI*

# METHODS OF APPROXIMATION

## 16-1 Approximation by Differentials

In certain kinds of work we find ourselves interested in how the values of a function vary when $x$ varies only a small amount away from some particular value $x_0$. For various reasons, usually in order to make things manageably simple, we may be satisfied with an *approximation* of the way $f(x)$ changes. The simplest kind of approximation is that in which it is assumed that $f(x)$ changes linearly. If $f$ is differentiable at $x_0$, then the best we can do with linear approximation is to use the line tangent to the curve $y = f(x)$ at $x_0$ as an approximation of the curve itself. This kind of approximation is often called approximation by differentials. The connection between differentials and the tangent line was brought out in the very definition of differentials (see § 5-1).

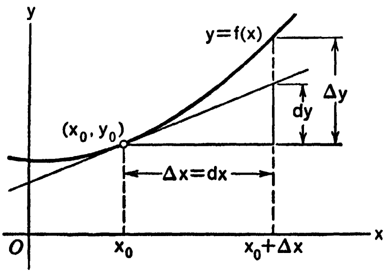



**Fig. 16-1**

It is helpful to see the whole situation graphically. We refer to Fig. 16-1. When $x$ changes from $x_0$ to $x_0 + \Delta x$, the true change is $\Delta y = f(x_0 + \Delta x) - f(x_0)$. But the differential approximation of $\Delta y$ is

$$dy = f'(x_0)\,dx, \quad \text{where} \quad dx = \Delta x. \tag{1}$$

Hence the approximation for the new value of $y$ is

$$y_0 + dy, \quad \text{or} \quad f(x_0) + f'(x_0)\,dx. \tag{2}$$

If we know how the graph appears we can judge something about the nature of our approximation. Evidently the true new value of $y$ is algebraically larger than the approximation if the curve lies above the tangent line, and the reverse is true if the curve lies below its tangent line.

**Example 1:** Use differential approximation to get a rough value of $\sqrt{142}$.

Since we know that $\sqrt{144} = 12$, we take $y = \sqrt{x}$, $x_0 = 144$, $y_0 = 12$, and $dx = -2$. Then calculating $dy$ and putting in the particular values, we get

$$dy = \frac{dx}{2\sqrt{x}} = \frac{-2}{2(12)} = -\frac{1}{12}.$$

Hence our approximation of the value of $\sqrt{142}$ is $12 - \frac{1}{12} = \frac{143}{12}$, or about 11.92.

**Example 2:** Calculate sin 46°, approximately.

In solving, we use radian measure, since our differentiation formulas presuppose that. We take $y = \sin x$. Since we know the sine of 45°, we take the radian equivalent, $\pi/4$, as our $x_0$. Then $dx$ must be the radian equivalent of 1°, which is $\pi/180$. Thus $y_0 = \sin(\pi/4) = \sqrt{2}/2$,

$$dy = \cos x\, dx = \frac{\sqrt{2}}{2}\frac{\pi}{180}.$$

Our approximation is thus

$$\sin 46° = \frac{\sqrt{2}}{2} + \frac{\pi\sqrt{2}}{360} = 0.7194.$$

A four-place table gives $\sin 46° = 0.7193$. Thus the method is reasonably accurate in this case.

The use of differential approximation is really the same as using Taylor's formula, going only to the term $n = 1$, and ignoring $R_1(x)$. Thus Taylor's formula for $f(x) = \sqrt{x}$ with $a = x_0$ and $n = 1$ is

$$\sqrt{x} = \sqrt{x_0} + \frac{1}{\sqrt{2x_0}}(x - x_0) + R_1(x).$$

If we put $x_0 = 144$, $x = 142$, and ignore $R_1(x)$, we get the same results as in Example 1.

*Percentage Error*

Sometimes we are concerned with errors in $y$ caused by using erroneous values of $x$. The errors may be due simply to faulty observation or to the natural limitations on precise physical measurements. If a true value of $x$ is measured as $x_0$, with the error represented by $dx$, then the difference between the true value $y = f(x_0 + dx)$ and the calculated value $y_0 = f(x_0)$ is approximately $dy = f'(x_0)\,dx$, and the approximate *relative error* in $y$ is $dy/y_0$. If this fraction is reduced to a decimal and multiplied by 100, we get what is called the *approximate percentage error* in $y$. Problems on the limitation of error are often conveniently solved approximately with the use of differentials.

**Example 3:** From an assortment of steel balls it is desired to select all those whose diameter is 1 centimeter. If the permissible percentage deviation in the diameters of the balls is 3%, and they are to be selected by weighing them, what is the approximate permissible deviation in their weights?

We take the density of steel as 7.6 grams per cubic centimeter. The weight of a steel ball $2r$ centimeters in diameter would be, in grams,

$$M = (7.6)\frac{4}{3}\pi r^3.$$

The approximate change in weight, due to a change $dr$ in the radius, would be

$$dM = (7.6)4\pi r^2\,dr.$$

Now, with $2r = 1$ centimeter as the standard diameter, a deviation of not over 3% in this diameter means that $\left|\frac{dr}{r}\right| \le 0.03$. Hence, when $r = 0.5$,

$$|dM| = \left|(7.6)4\pi r^3\frac{dr}{r}\right| \le (7.6)(4\pi)(0.5)^3(0.03),$$

or,

$$|dM| \le 0.358.$$

This means that the permissible deviation in weight, with the weight of a ball 1 centimeter in diameter as a standard, is about 0.36 gram.

**Example 4:** What is the approximate permissible percentage deviation in $M$ in Example 4?

Since we wish to find $dM/M$, it is convenient to begin by taking the logarithms of both members of the formula for $M$:

$$\log M = \log(7.6) + \log\frac{4\pi}{3} + 3\log r.$$

Then

$$\frac{1}{M}\frac{dM}{dr} = \frac{3}{r}, \quad \text{or} \quad \frac{dM}{M} = \frac{3\,dr}{r}.$$

As before,

$$\left|\frac{dr}{r}\right| \le 0.03, \quad \text{and so} \quad \left|\frac{dM}{M}\right| \le 0.09.$$

The approximate permissible percentage deviation in $M$ is therefore 9%.

## EXERCISES

**1.** Compute approximately, by differentials: (a) $\sqrt{37}$; (b) $\sqrt{50}$; (c) $\sqrt{98.5}$; (d) $\sqrt[3]{121}$; (e) $\sqrt[5]{30}$; (f) $(201)^4 - (200)^4$.

**2.** Compute approximately, by differentials: (a) tan 44°; (b) cos 59°; (c) csc 31°.

**3.** From $\log 5 = 1.6094$ estimate $\log 5.15$ by differentials. Is the result too small or too large? Why?

4. If two spheres have surface areas $S_1$, $S_2$, respectively, and the radii are each increased by the same slight amount, show that the increases in the volumes are approximately in the ratio $S_1:S_2$.

5. A cubical wooden block originally had edges 24 inches long. Then a layer $\frac{1}{8}$ inch thick was cut off of each face. Find (a) the approximate decrease in volume of the block; (b) the approximate decrease in surface area of the block; (c) the percentage decrease in edge length, and the approximate percentage decreases in surface area and volume, respectively.

6. Let $f(x) = x^2$. Interpret $f(x)$ as the area of a square of side $x$. By drawing this square and also a square of side $x + \Delta x$, show diagramatically the quantities $\Delta y$ and $dy$ if $dx = \Delta x$. Show the geometrical representation of $\Delta y - dy$ as an area.

7. Carry out for $f(x) = x^3$ a project like that in Exercise 6, using cubes instead of squares.

8. The width of a river is calculated by measuring the angle of elevation, from a point on one bank, of a tree 80 feet high on the opposite bank. If the angle of elevation is observed to be 30°, with a possible error of 10′, estimate the width of the river and display an approximate limitation on the amount by which the estimate may be wrong.

9. The height of a flagpole is to be calculated by measuring the length of its shadow and at the same time observing the angle of elevation of the top of the flagpole from the end of the shadow. If the calculated height is to be correct to within 1%, and the shadow is measured perfectly accurately, show that, for an angle of elevation of approximately 60°, the permissible error in measuring this angle cannot exceed about 15′.

10. A plank $2b$ feet long is laid across a cylindrical pipe $a$ feet in diameter, the plank balancing in a horizontal position with its mid-point in contact with the top of the pipe. If the plank is disturbed slightly and turns through an angle $\theta$ (without slipping on the pipe), one end goes down a distance $s$ feet below its original level. Show that $s = a(1 - \cos\theta) + (b - a\theta)\sin\theta$. If $\theta$ is quite small, obtain a simpler approximate formula for $s$.

## 16-2 The Intersection of Two Curves

Sometimes we wish to find where two curves intersect, but the equations of the curves are such that a neat exact solution is not available. We then have to resort to some method of getting the solution approximately. A good way of explaining procedures for such problems is to demonstrate what can be done in a particular case.

**Example 1:** Consider the curve $y = \sin^2 x$, $0 \leq x \leq \pi$. It is required to find the line $y = mx$ which is tangent to the given curve at a certain point $x_0$

for which $0 < x_0 < \pi/2$ (see Fig. 16-2). The figure itself, if drawn with care, suggests something like $x_0 = 1.12$ as an approximate answer. But let us formulate the problem in terms of equations. The slope of the curve at $x_0$ is $2 \sin x_0 \cos x_0$ (as we see by computing $dy/dx$). Hence $m$ and $x_0$ must satisfy the equations

$$mx_0 = \sin^2 x_0, \qquad m = 2 \sin x_0 \cos x_0.$$

If we eliminate $m$ by division, we find

$$x_0 = \frac{\sin^2 x_0}{2 \sin x_0 \cos x_0} = \frac{1}{2} \tan x_0,$$

or $2x_0 = \tan x_0$. It is then convenient to think of $x_0$ as a number determined by the intersection of the two curves

$$y = 2x, \qquad y = \tan x, \tag{1}$$

with the proviso that $0 < x_0 < \pi/2$. We can easily draw good graphs of the two curves in (1), and this gives us a starting point for more refined methods

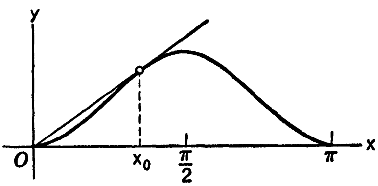



Fig. 16-2

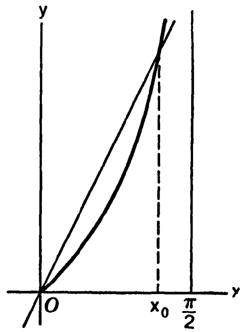



Fig. 16-3

of calculating $x_0$ approximately. See Fig. 16-3. In this case we can use tables to good advantage. Using our rough estimate $x = 1.12$, we look in Table III, comparing the entries for $x$ and $\tan x$. We want to have $2x = \tan x$, and we soon notice that the $x$ we want is between 1.16 and 1.17, because $2x - \tan x$ changes from positive to negative here:

| $x$ | $2x$ | $\tan x$ | $2x - \tan x$ |
|---|---|---|---|
| 1.16 | 2.32 | 2.2958 | 0.0242 |
| 1.17 | 2.34 | 2.3600 | −0.0200 |

With this table the best we can do now is make an interpolation. To reduce $2x - \tan x$ from 0.0242 to zero is to go the fractional part $242/442 = 0.55$ of the way from 0.0242 to −0.0200. Hence we estimate that

$$x_0 = 1.16 + 0.55(1.17 - 1.16) = 1.1655.$$

Alternatively, we could have started out using Table IV, observing that $x_0$ must be between 1.1636 and 1.1665. A similar interpolation in this case gives

$x_0 = 1.1656$. We leave the problem at this point. As for the slope of the line in Fig. 16-2, it is $m = 2 \sin x_0 \cos x_0 = \sin 2x_0$, which with $x_0 = 1.1655$ becomes $m = \sin 2.3310$. Since our tables do not extend this far we use $\sin A = \sin (\pi - A)$, so that $m = \sin (3.1416 - 2.3310) = \sin (0.8106) = 0.7248$, approximately, by interpolation in Table IV.

### *A Method of Successive Approximations*

Next we illustrate a method which has a certain theoretical interest and which is sometimes expedient in practice.

**Example 2:** Find the intersection of the parabola $(y - 3)^2 = -12(x - 3)$ and the hyperbola $y(4 - x) = 2$.

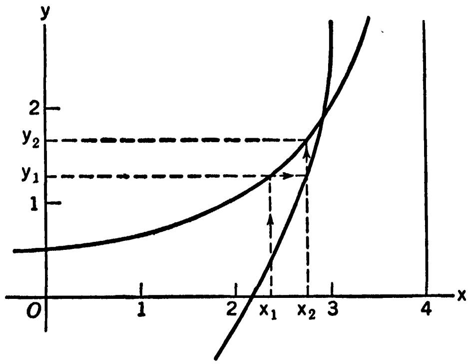



**Fig. 16-4**

In Fig. 16-4 we have shown the graphs in the vicinity of the point of intersection. The hyperbola has asymptotes $x = 4$, $y = 0$, and the parabola opens to the left, with vertex at (3, 3). In this case we see that the intersection comes at a point where $x$ is a bit less than 3 and $y$ is a bit less than 2. We start with an estimated value $x_1$, use the equation of the hyperbola to compute $y_1$ from this $x_1$, then use the equation of the parabola to compute a second $x$-value $x_2$, and repeat the process. If the equation of the hyperbola is written as $y = f(x)$, while the parabola is written as $x = g(y)$, the procedure goes as follows:

$$y_1 = f(x_1), \qquad x_2 = g(y_1)$$

$$y_2 = f(x_2), \qquad x_3 = g(y_2)$$

and so on. It is evident from the diagram that the sequence $\{x_n\}$ converges to a limit which is the $x$-coordinate of the point of intersection. If we take $x_1$ too large, the sequence $\{x_n\}$ will decrease toward its limit, instead of increasing as in the diagram.

It is important to calculate $y_n$ from $x_n$, using the correct one of the two curves. If we were to use the parabola instead of the hyperbola in this case, things would get worse instead of better. If the two curves have slopes of opposite signs, the sequence $\{x_n\}$ will not be monotonic, but the method may still be used. A graph should always be used to guide the work correctly. If the curves cross more or less like lines of slopes $+1$ and $-1$, respectively, the successive approximations may not improve. Finally, for practical effectiveness, we must be able to compute $y_n$ and $x_{n+1}$ successively with reasonable ease.

We now list a few calculations, starting from $x_1 = 2.8$. The $x_1$ and $x_2$ in Fig. 16-4 are shown as much less than we would reasonably guess, merely to show things up clearly in the diagram.

$$y = f(x) = \frac{2}{4 - x}, \qquad x = g(y) = 3 - \tfrac{1}{12}(3 - y)^2,$$

$$\begin{aligned} &x_1 = 2.8, \quad y_1 = 1.67, && y_1 = 1.67, x_2 = 2.85,\\ &x_2 = 2.85, \quad y_2 = 1.74, && y_2 = 1.74, x_3 = 2.87,\\ &x_3 = 2.87, \quad y_3 = 1.77, && y_3 = 1.77, x_4 = 2.874,\\ &x_4 = 2.874, y_4 = 1.776. \end{aligned}$$

In the next section we consider a different procedure, that of Newton's method, which is generally preferred over the method we have just discussed. But no one method for attacking problems of this kind can fairly be said to be the best in all situations.

## EXERCISES

1. Find where $y = x$ intersects $y = \cos x$, using a graph, Table III, and interpolation.

2. Find the intersection of $y = e^x - 1$ and $y = \log(1/x)$ using a graph, Tables I and II, and interpolation.

3. Find $x$ such that $0 < x < \pi/2$ and $e^x = \tan x$, using a graph, Tables II and III, and interpolation. Compare your result with what you get by starting with $x_1 = 1.2$ and computing $x_2$ by the successive approximation method.

4. Draw graphs of $y = 1 - (x^2/4)$ and $x = \log(1 + y)$ well enough to show approximately where they intersect in the first quadrant. Then, starting with $x_1 = 0.6$, calculate $x_2$ and $x_3$ by the method of successive approximations. Use Table I.

5. If, in the previous exercise, $x = \log(1 + y)$ is written $y = e^x - 1$, use Table II and interpolation to get a solution of $e^x - 1 = 1 - (x^2/4)$.

6. Show that arc length from $(0, a)$ to $(x, y)$ along $y = a \cosh(x/a)$ (where $a > 0$) is $s = a \sinh(x/a)$. If $s$ and $x$ are assigned certain positive values, the foregoing equation becomes an equation from which to solve for $a$. It is more convenient to let $t = x/a$, and then $t$ is to be found from $\sinh t = st/x$. (a) Solve for $t$ if $x = 100$ and $s = 120$, using Table II. (b) The curve $y = a \cosh(x/a)$ is the curve in which a long rope hangs if stretched not quite tight between two points. Using the result in (a), find how much the mid-point of a 240 foot rope is below the level of its ends if these ends are on the same horizontal line and 200 feet apart.

## 16-3 Newton's Method

The problem of finding where the line $y = 2x$ intersects the curve $y = \tan x$ (see Example 1, § 16-2) is the same as the problem of solving the equation

$$2x - \tan x = 0. \tag{1}$$

Likewise, the problem in Example 2 of § 16-2 can be restated as a problem of solving a single equation. If we eliminate $y$ between the equations of the hyperbola and parabola, we obtain the cubic equation

$$(3x - 10)^2 = -12(x - 3)(x - 4)^2,$$

or

$$12x^3 - 123x^2 + 420x - 476 = 0. \tag{2}$$

These are examples of equations of the form $f(x) = 0$ which we cannot solve by simple formulas. As a practical matter, we must be satisfied with approximate solutions. The method we now consider, known as Newton's method, is based on a simple geometric idea. Let $x_1$ be an approximation of a value of $x$ such that $f(x) = 0$. Let $y_1 = f(x_1)$ and draw the tangent to the curve $y = f(x)$ at $(x_1, y_1)$ (see Fig. 16-5). Let $x_2$ be the abscissa of the point where the tangent crosses the $x$-axis. Then $x_2$ may be used as a second approximation to the solution. The formula for finding $x_2$ is easily worked out. The equation of the tangent is

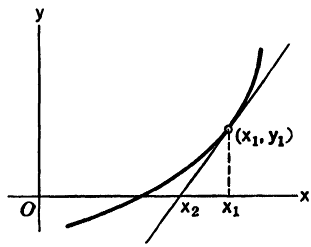


**Fig. 16-5**

$$y - y_1 = f'(x_1)(x - x_1).$$

When it crosses the $x$-axis we have $y = 0$, $x = x_2$, so that

$$-y_1 = f'(x_1)(x_2 - x_1),$$

or

$$x_2 = x_1 - \frac{y_1}{f'(x_1)} = x_1 - \frac{f(x_1)}{f'(x_1)}. \tag{3}$$

The process may be repeated, giving

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n = 1, 2, \cdots. \tag{4}$$

There are, of course, certain conditions which must be observed in applying this procedure. If no restrictions are imposed, the process may not yield a sequence $\{x_n\}$ which does anything useful. This is illustrated in Exercise 1. In the next paragraph we describe some reasonable conditions which will guarantee a useful outcome of applying Newton's method.

Suppose that $f$ has first and second derivatives when $a \leq x \leq b$, that $f(a)$ and $f(b)$ are of opposite sign, and that $f'(x)$ and $f''(x)$ are each of constant sign when $a \leq x \leq b$, so that there are no horizontal tangents and no points of inflection corresponding to the $x$'s under consideration. Then, as $x$ goes from $a$ to $b$, the value $y = f(x)$ is either always increasing or always decreasing, and hence (since it changes sign) there is a unique $x$ between $a$ and $b$ for which $f(x) = 0$. Let $\bar{x}$ denote this root of the equation. Suppose the starting value $x_1$ in Newton's method is such that $a \leq x_1 \leq b$ and that $f(x_1)$ has the same sign as that of $f''(x)$ [i.e., $f(x_1) > 0$ if the curve is concave upward, $f(x_1) < 0$ if the curve is concave downward]. Then it is possible to prove that the sequence $\{x_n\}$ converges to $\bar{x}$ and that each $x_n$ is closer to $\bar{x}$ than its predecessor $x_{n-1}$. In practice, a very few applications of the method are sufficient to give the root accurately to several decimal places.

If the starting value $x_1$ is such that $f(x_1)$ has sign opposite to that of $f''(x)$, the root $\bar{x}$ will be between $x_1$ and $x_2$. It can happen in this case that $x_2$ is not so close to $\bar{x}$ as $x_1$ is, and that $x_3$ is even worse. But if $x_2$ is also in the interval $a \leq x \leq b$, the succeeding approximations will steadily improve and converge to $\bar{x}$.

A valuable indication of the rapidity of convergence of the approximations in Newton's method is given by the following statement, whose proof we omit: Suppose, along with what has already been said, that $f''(x)$ is continuous. Let $M$ be the maximum of $|f''(x)|$ and $m$ the minimum of $|f'(x)|$ on the interval $[a, b]$. Finally, suppose $M \leq 4m$. Then, assuming that $x_1$ and $x_2$ both lie in the interval, it can be shown that if any particular $x_k$ approximates $\bar{x}$ with accuracy to a certain number of decimal places, then $x_{k+1}$ has accuracy to twice as many decimal places.

**Example:** Find the point on the parabola $2y = x^2$ which is nearest the point $(2, 0)$.

If $(x, y)$ is on the parabola, then $y = x^2/2$ and the square of the distance from $(x, y)$ to $(2, 0)$ is

$$D^2 = (x - 2)^2 + \frac{x^4}{4}.$$

We want the value of $x$ which makes $D^2$ as small as possible. The condition for this minimum is that

$$\frac{d}{dx}(D^2) = 2(x - 2) + x^3 = 0.$$

Hence we must solve the equation $x^3 + 2x - 4 = 0$. A rough sketch of the parabola shows us that we must expect $x$ to be slightly larger than 1. If we set $y = f(x) = x^3 + 2x - 4$, we find

$$f'(x) = 3x^2 + 2, \qquad f''(x) = 6x.$$

Thus $f'(x)$ is always positive and $f''(x) > 0$ when $0 < x$. Calculation shows that $f(1) = -1$, $f(2) = 8$, and linear interpolation gives us $x = 1 + \frac{1}{9} = 1.11\cdots$ as an estimate of the root we are seeking. This estimate is too small, for the curve is concave upward, and hence the chord from $(1, -1)$ to $(2, 8)$

cuts the $x$-axis to the left of the point where the curve crosses the axis. The student should make a graph and visualize what we are saying here.

Before starting in with Newton's method let us attempt to get a good one-decimal place estimate of the root. We try $x = 1.15$ and find $f(x) = -0.179$, which shows that the root is larger than 1.15. We try $x = 1.2$ and find $f(x) = 0.128$, which shows that the root is less than 1.2. Hence, with one-place accuracy, the root is 1.2. The calculations of $f(x)$ in these cases can be made conveniently by synthetic division.

Now we shall carry out Newton's method, starting with $x_1 = 1.2$. Then

$$f(1.2) = 0.128, \qquad f'(1.2) = 6.32,$$

$$x_2 = 1.2 - \frac{0.128}{6.32} = 1.18.$$

We are justified in claiming two-place accuracy for this result, for, if we use the interval $[1, 2]$, $M$ = maximum of $|f''(x)| = 12$ and $m$ = minimum of $|f'(x)| = 5$, and so $M \leq 4m$ in this case. The next approximation will give us four-place accuracy:

$$f(1.18) = 0.00303, \qquad f'(1.18) = 6.1772$$

$$x_3 = 1.18 - \frac{0.00303}{6.1772} = 1.1795.$$

In ordinary practice the $M \leq 4m$ test is not always used, since it may be laborious to check whether it is fulfilled. A common procedure is to stop the approximations as soon as two successive ones agree to the required number of places.

## EXERCISES

1. The polynomial $f(x) = 24x^3 - 18x^2 + 1$ has three roots, one of which is between 0 and $\frac{1}{2}$. If we start Newton's method with $x_1 = 0.2$ and compute $x_2, x_3, \cdots$, does the sequence $\{x_n\}$ converge to the root between 0 and $\frac{1}{2}$? What does it do? Base your discussion on a carefully constructed graph.

2. (a) Calculate to two decimal places the root of $x^3 - 3x^2 + 3 = 0$ which is between 2 and 3. (b) Obtain the root to four decimal places.

3. Find the other roots of $x^3 - 3x^2 + 3 = 0$, each to two decimal places.

4. Find the abscissa of the point of intersection of the curves $y = x^3$, $y = 2 - 2x$. Begin by locating the root between consecutive tenths, and make the choice of $x_1$ to two decimal places by linear interpolation. Then calculate $x_2$.

5. A spherical ball of radius 2 inches and specific gravity $\frac{1}{4}$ floats on water. Show that the depth $x$ to which the ball is submerged is a root of the equation $x^3 - 6x^2 + 8 = 0$. Graph the function $y = x^3 - 6x^2 + 8$ in the in-

terval $0 \leq x \leq 4$, and obtain a first approximation to the desired root by assuming that the graph is a straight line between the points for which $x = 1$ and $x = 2$. Then obtain a second approximation, using Newton's method.

6. The equation $hx^3 - x^2 + x + 2 = 0$ has a root $x = -1$ if $h = 0$. It therefore has a root near $-1$ if $h$ is small. Show by Newton's method that

$$x = -1 + \frac{1}{3}\left(\frac{h}{h+1}\right)$$

is a good approximation to this root.

7. Find the positive root of the equation $e^{-x} + \frac{1}{2}x - 1 = 0$. First graph $y = e^{-x}$ and $y = 1 - \frac{1}{2}x$ on the same axes, and estimate the roots of the original equation by finding where the two graphs intersect. Then determine the positive root more accurately by Newton's method.

8. In each of the following problems solve for $x$ (subject to the stated restriction in some cases). Get your initial estimate of the solution by graphing two curves well enough to get a fair idea of where they intersect.
   (a) $x = 5 \log x$, $1.2 < x < 1.3$.
   (b) $e^x = 2 - x$.
   (c) $\cos x = 10x$.
   (d) $x^2 = 2 \cos x$, $x > 0$.

9. If a rope hangs over a rough circular cylinder of radius $r$ whose axis is horizontal, the rope will barely be held in place by friction if one end is at the level of the axis of the cylinder and the other end hangs down a distance $L$ below the level of the axis on the other side, where $(1 + \mu^2)L = 2r\mu(1 + e^{\pi\mu})$, $\mu$ being the coefficient of friction. Find the value of $\mu$ in this situation if $L = \pi r$. Suggestion: let $\pi\mu = x$ and solve for $x$. Begin by getting a reasonably good one-decimal place estimate of $x$, and then improve it by Newton's method.

## 16-4 Approximating Definite Integrals

In this section we shall consider two methods of computing the value of an integral

$$\int_a^b f(x)\, dx \tag{1}$$

approximately, by formulas which employ the values of $f(x)$ at a finite number of points on the interval $[a, b]$. By comparison with the use of the approximating sums which are used in defining the integral (see § 6-1), the formulas of this section usually give better approximations for the same amount of labor devoted to computation.

### *The Trapezoidal Rule*

Suppose that $[a, b]$ is divided into $n$ equal parts by points $x_0, x_1, \cdots, x_n$ in order from $x = a$ to $x = b$. Let $y_0, \cdots, y_n$ be the corresponding values of $y = f(x)$. We then approximate the area between $y = f(x)$ and the $x$-axis, for $x_{k-1} \le x \le x_k$, by means of a trapezoid whose oblique side joins the points $(x_{k-1}, y_{k-1})$ and $(x_k, y_k)$ (see Fig. 16-6). The area of this trapezoid is

$$\tfrac{1}{2}(y_{k-1} + y_k)(x_k - x_{k-1}). \qquad (2)$$

If we write
$$x_k - x_{k-1} = \Delta x = \frac{b - a}{n}, \qquad (3)$$

the addition of the expressions (2) for $k = 1, 2, \cdots, n$ yields the sum

$$(\tfrac{1}{2}y_0 + y_1 + y_2 + \cdots + y_{n-1} + \tfrac{1}{2}y_n)\,\Delta x \qquad (4)$$

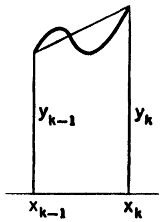


**Fig. 16-6**

as an approximate value of the integral (1). This is called approximation by the trapezoidal rule.

**Example 1:** Use the trapezoidal rule with $n = 4$ to get an approximate value of

$$\int_0^1 (1 + x^3)^{1/2}\,dx.$$

Here $f(x) = (1 + x^3)^{1/2}$ and $x_0 = 0$, $x_2 = \frac{1}{2}$, $x_1 = \frac{1}{4}$, $x_3 = \frac{3}{4}$, $x_4 = 1$. We can compute the $y$'s easily from a table of square roots:

$$y_0 = (1)^{1/2} = 1.000,$$
$$y_1 = \left(\frac{65}{64}\right)^{1/2} = 1.008,$$
$$y_2 = \left(\frac{9}{8}\right)^{1/2} = 1.061,$$
$$y_3 = \left(\frac{91}{64}\right)^{1/2} = 1.192,$$
$$y_4 = (2)^{1/2} = 1.414.$$

Thus, approximately,

$$\int_0^1 (1 + x^3)^{1/2}\,dx = (0.500 + 1.008 + 1.061 + 1.192 + 0.707)\,\frac{1}{4} = 1.117.$$

*Simpson's Rule*

This method is based on a more ingenious device than the use of trapezoids. For Simpson's rule we again divide $[a, b]$ into $n$ equal parts, but we insist that $n$ be an *even* integer. Now consider the first three points $x_0$, $x_1$, $x_2$, and the corresponding points on the curve $y = f(x)$. If these points are not collinear there is a unique parabola with its axis parallel to the $y$-axis, the parabola passing through the three points. The equation of a parabola with its axis parallel to the axis is of the form $y = P(x)$, where $P(x)$ is a quadratic polynomial, and we may write $P(x)$ in the form

$$P(x) = A + B(x - x_1) + C(x - x_1)^2, \qquad (5)$$

by choosing $A$, $B$, and $C$ suitably. We shall choose them so as to make the three points under consideration lie on the parabola. The conditions are

$$x = x_0, \qquad A + B(x_0 - x_1) + C(x_0 - x_1)^2 = y_0, \tag{6}$$

$$x = x_1, \qquad A = y_1,$$

$$x = x_2, \qquad A + B(x_2 - x_1) + C(x_2 - x_1)^2 = y_2. \tag{7}$$

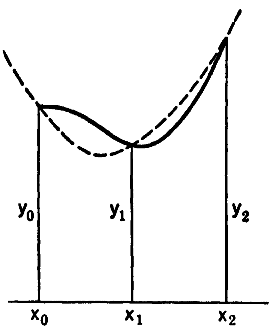



**Fig. 16-7**

Equations (6) and (7) can be used to solve for $B$ and $C$. It is more convenient to write them in the form

$$B\,\Delta x + C(\Delta x)^2 = y_2 - y_1,$$

$$-B\,\Delta x + C(\Delta x)^2 = y_0 - y_1,$$

by making use of the definition of $\Delta x$ and the fact that $A = y_1$. In particular, note that

$$2C(\Delta x)^2 = y_0 - 2y_1 + y_2. \tag{8}$$

For a diagram of the parabola and the three points, see Fig. 16-7. The parabola is shown dotted; the other curve is $y = f(x)$.

Now we shall think of the parabola as an approximation to the curve $y = f(x)$ in the interval from $x_0$ to $x_2$, and compute this part of the integral accordingly. Thus we obtain the approximation

$$\int_{x_0}^{x_2} f(x)\,dx = \int_{x_0}^{x_2} [A + B(x - x_1) + C(x - x_1)^2]\,dx$$

$$= \left[Ax + \frac{1}{2}B(x - x_1)^2 + \frac{1}{3}C(x - x_1)^3\right]_{x_0}^{x_2}.$$

On evaluating this and recalling the definition of $\Delta x$, we obtain the expression

$$2A\,\Delta x + \tfrac{2}{3}C(\Delta x)^3.$$

By using the values found for $A$ and $C$, we put this in the form

$$2y_1\,\Delta x + \tfrac{1}{3}(y_0 - 2y_1 + y_2)\,\Delta x = \tfrac{1}{3}(y_0 + 4y_1 + y_2)\,\Delta x.$$

We can do the same sort of thing with the intervals $[x_2, x_4]$, $[x_4, x_6]$, $\cdots$. When the results are all added together we get the approximation formula

$$\int_a^b f(x)\,dx = \frac{\Delta x}{3}(y_0 + 4y_1 + 2y_2 + \cdots + 4y_{n-1} + y_n). \tag{9}$$

This is known as Simpson's rule. Notice the arrangement of the terms in the parentheses: $y_0$ and $y_n$ occur with factor 1, the remaining $y$'s with even subscripts occur with factor 2, and the $y$'s with odd subscripts occur with factor 4.

**Example 2:** Use Simpson's rule with $n = 4$ to calculate

$$\int_0^2 \frac{dx}{1 + x^4}.$$

The tabulations are

$$\begin{aligned} y_0 &= 1.000, & y_0 &= 1.000 \\ y_1 &= 16/17 = 0.941, & 4y_1 &= 3.764 \\ y_2 &= 1/2 \quad = 0.500, & 2y_2 &= 1.000 \\ y_3 &= 16/97 = 0.165, & 4y_3 &= 0.660 \\ y_4 &= 1/17 \quad = 0.059, & \underline{y_4} &= \underline{0.059} \\ & & & \phantom{=}\ 6.483 \end{aligned}$$

Hence, approximately,

$$\int_0^2 \frac{dx}{1 + x^4} = \frac{1}{6}(6.483) = 1.080.$$

## EXERCISES

1. The exact value of $\int_0^1 \frac{dx}{1 + x^2}$ is, of course, $\pi/4$. Using $n = 6$, calculate the integral approximately (a) by the trapezoidal rule, and (b) by Simpson's rule. Carry the work to three decimal places and round off to two places in the final result.

2. Approximate the value of $\int_0^1 e^{-x^2}\, dx$ with $n = 4$, (a) by the trapezoidal rule, (b) by Simpson's rule, (c) by using an infinite series. Give answers to three decimal places. The series may be used to give assured three-place accuracy. In (a) and (b) the calculations can be made from Table II.

3. The arc of the curve $y = \log x$ from $x = 1$ to $x = 2$ is revolved about the $x$-axis. Express as an integral the area of the resulting surface of revolution, and calculate its value approximately by Simpson's rule with $n = 4$.

4. (a) What does Simpson's rule with $n = 4$ give for the problem of Example 1? (b) By examining the concavity of the curve prove that the answer 1.117 given by the trapezoidal rule is certainly too large. (c) Use a binomial series and then integrate it to get an approximate value of the integral which is correct to three places of decimals.

5. In the derivation of Simpson's rule, if the three points on the curve, corresponding to $x_0$, $x_1$, $x_2$, are collinear, then instead of having a parabola through them we get a straight line through them. Check this by examining (5) and (8) and explaining what you find.

6. If $n = 2$, Simpson's rule becomes

$$\int_a^b f(x)\, dx = \frac{b - a}{6}\left[f(a) + 4f\left(\frac{a + b}{2}\right) + f(b)\right].$$

This can be used to obtain an approximate formula for the volume of a solid:

$$V = \frac{h}{6}(B_1 + 4M + B_2),$$

where $B_1$, $B_2$, and $M$ refer to the areas of plane sections of the solid, all perpendicular to a single axis. The end-section areas are $B_1$ and $B_2$, and $M$ is the area of a section halfway between, while $h$ is the distance between the end sections. This formula for a volume is called the *prismoidal rule*. Explain why there is this connection between Simpson's rule and a volume formula.

**7.** Show that Simpson's rule gives an *exact* result with $f(x) = x^3$. Note that it is sufficient to prove this for $n = 2$. *Suggestion:* Let $\Delta x = h$, $b = a + 2h$, and express everything in terms of $a$ and $h$. Now explain why Simpson's rule gives an exact result for $y = P(x)$, where $P(x)$ is any polynomial of degree 3 or less.

**8.** Use the prismoidal rule (Exercise 6) to find the following volumes: (a) of a segment cut from a sphere of radius 5 by a diametral plane and a parallel plane 4 units from it; (b) of a frustum of a right circular cone if its end radii are 2 and 4, respectively, and its altitude is 6.

**9.** Compute $\int_0^\pi \frac{\sin x}{x}\,dx$ to three places of decimals by Simpson's rule with $n = 6$.

**10.** Find the length of the first quadrant arc of the ellipse $16x^2 + 25y^2 = 400$, by using the parametrization $x = 5\sin t$, $y = 4\cos t$, and approximating the integral, (a) by Simpson's rule with $n = 2$; (b) by the trapezoidal rule with $n = 3$.

# CHAPTER XVII

# DETERMINANTS AND LINEAR SYSTEMS

## 17-1 Determinants of Order Two

Determinants of order two come to our attention naturally when we examine in general terms the problem of trying to solve a system of two equations in two unknowns. In order to see clearly the essential nature of this problem from our present point of view it is desirable to use a notation that allows us to realize fully the algebraic symmetry which is involved. Let us write our two equations in the form

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned}\right\} \quad (1)$$

The subscripts serve simply to distinguish one literal quantity from another. The $a$'s and $b$'s are given, and we wish to solve for the $x$'s.

If we multiply the first equation by $a_{22}$, the second by $-a_{12}$, and add, we obtain

$$(a_{11}a_{22} - a_{21}a_{12})x_1 = b_1a_{22} - b_2a_{12}. \quad (2)$$

Likewise, multiplying the first equation by $-a_{21}$, the second by $a_{11}$, and adding, we obtain

$$(a_{11}a_{22} - a_{21}a_{12})x_2 = a_{11}b_2 - a_{21}b_1. \quad (3)$$

From these considerations we conclude that if $x_1$ and $x_2$ satisfy equations (1), then they also satisfy equations (2) and (3). For convenience let us write

$$D = a_{11}a_{22} - a_{21}a_{12}. \quad (4)$$

Now suppose that $x_1$ and $x_2$ satisfy (2) and (3). Will they then satisfy the system (1)? To investigate this, multiply (2) through by $a_{11}$ and (3) by $a_{12}$; then add. The result is

$$D(a_{11}x_1 + a_{12}x_2) = a_{11}(b_1a_{22} - b_2a_{12}) + a_{12}(a_{11}b_2 - a_{21}b_1).$$

On simplifying the right side we see that

$$D(a_{11}x_1 + a_{12}x_2) = Db_1. \tag{5}$$

In a similar way we see that if (2) and (3) hold, then

$$D(a_{21}x_1 + a_{22}x_2) = Db_2. \tag{6}$$

Now, (5) and (6) together are equivalent to (1) if $D \neq 0$, for then we can cancel $D$ in (5) and (6). Hence we can say: *If $D \neq 0$, then equations* (1) *have a uniquely determined solution for $x_1$ and $x_2$, namely,*

$$x_1 = \frac{b_1a_{22} - b_2a_{21}}{D}, \qquad x_2 = \frac{a_{11}b_2 - a_{21}b_1}{D}. \tag{7}$$

This is what is called Cramer's rule for a system of two linear equations in two unknowns.

For the present we put off the consideration of what can be said if $D = 0$.

Our concern here is not with the practical problem of solving simultaneous linear equations. Instead, we are interested in the appearance of formulas (2) and (3). There is obviously a recurring pattern here—a pattern which involves four quantities $a_{11}$, $a_{12}$, $a_{21}$, $a_{22}$. In functional notation we might write

$$D = F(a_{11}, a_{12}, a_{21}, a_{22}). \tag{8}$$

Then $D$ is the value of the function $F$ when values are assigned to the $a$'s. This same function can be considered with other symbols for the variables. Comparing (4) and (8) we see that

$$F(u, v, x, y) = uy - xv. \tag{9}$$

In order to make it easier to remember the order in which the variables occur, it is convenient to use the schematic arrangement

$$F(u, v, x, y) = \begin{vmatrix} u & v \\ x & y \end{vmatrix}, \tag{10}$$

in which the first two variables are written in the first row, the second two in the second row. Comparing (9) and (10), we observe that $uy$ is the product of terms on one diagonal of the square array in (10), while $xv$ is the product of terms on the other diagonal.

This particular function of four variables is called *a determinant of order two*. The individual numbers $u$, $v$, $x$, $y$ are called *entries* of the determinant.

When we exhibit a particular value of the function in the schematic array (10), we often refer to it as a determinant, although it is, strictly speaking, *a value of the determinant function.*

**Example 1:**

$$F(3, 5, -1, 0) = \begin{vmatrix} 3 & 5 \\ -1 & 0 \end{vmatrix} = 3(0) - (-1)(5) = 5.$$

In the determinant notation we can now exhibit equations (2) and (3) in the form

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} x_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \tag{11}$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} x_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}. \tag{12}$$

The algebraic symmetry of these equations is noteworthy. Notice in particular the appearance of the determinant on the left in relation to the arrangement of coefficients in the original system (1). Note also how the determinants on the right are obtained from the one on the left. In one case $b_1$ and $b_2$ displace the first column, and in the other case they displace the second column.

Now let us inquire, when is a value of the determinant function zero? There are two cases: (1) The value is zero if there is either a row or a column in which the entries are both zero. (2) If we do not have case (1), then the value of the determinant is zero if and only if the entries in the second row are proportional to the entries in the first row; that is, if and only if

$$\frac{a_{11}}{a_{21}} = \frac{a_{12}}{a_{22}}.$$

An alternative way of stating it is that the entries in the second column are to be proportional to those in the first column.

A simple way of avoiding the separation into two cases is available if we explain what is meant by saying that one number pair $(u, v)$ is a multiple of another pair $(x, y)$. To say that $(u, v) = k\cdot(x, y)$ means that $(u, v) = (kx, ky)$, i.e., that $u = kx$ and $v = ky$. (We could, if we wished, call the number pairs vectors.) Then the value of the determinant

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

is 0 if and only if at least one of the rows is a multiple of the other row. It might be the zero multiple.

**Example 2:**

$$\begin{vmatrix} a & b \\ ka & kb \end{vmatrix} = a(kb) - (ka)b = k(ab - ab) = 0.$$

*Two Homogeneous Equations in Three Unknowns*

Consider two equations of the form

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0. \end{aligned}\right\} \quad (13)$$

These are called *homogeneous* because of the fact that if a triple of numbers $(x_1, x_2, x_3)$ satisfies both equations, so does any multiple of the triple, such as $k(x_1, x_2, x_3) = (kx_1, kx_2, kx_3)$. If one or both zeros on the right side in (13) were replaced by nonzero constants, the system would no longer be homogeneous.

We shall prove the following theorem:

THEOREM 17-A. *Suppose that at least one of the three determinants*

$$c_1 = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}, \qquad c_2 = \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix}, \qquad c_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \qquad (14)$$

*is not zero. Then all of the solutions* $(x_1, x_2, x_3)$ *of the system* (13) *are given by forming multiples of the triple* $(c_1, c_2, c_3)$. *That is,* $(c_1, c_2, c_3)$ *is a solution of* (13), *and every solution is a multiple of this particular solution.*

*Proof.* To prove that $x_1 = c_1$, $x_2 = c_2$, $x_3 = c_3$ is a solution of (13), we merely substitute. Upon calculation of $a_{11}c_1 + a_{12}c_2 + a_{13}c_3$, we find that it is zero. Likewise for the other equation. The student should work out the details. To go in the other direction, we assume that $(x_1, x_2, x_3)$ is a triple satisfying (13), and we must find a constant $k$ such that $x_1 = kc_1$, $x_2 = kc_2$, $x_3 = kc_3$. It is assumed that at least one of the $c$'s is not zero. Let us, for definiteness, assume that $c_3 \neq 0$. We can then apply what we have learned earlier, using (11) and (12) with $b_1 = -a_{13}x_3$, $b_2 = -a_{23}x_3$. (One must notice that this choice of the $b$'s makes (13) like (1) in form.) Thus, by (11),

$$c_3x_1 = \begin{vmatrix} -a_{13}x_3 & a_{12} \\ -a_{23}x_3 & a_{22} \end{vmatrix} = -x_3a_{13}a_{22} + x_3a_{23}a_{12}$$

$$= x_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = c_1x_3.$$

Likewise, using (12) we find that $c_3x_2 = c_2x_3$. We now choose $k = x_3/c_3$, so that $x_3 = kc_3$. Then $c_3x_1 = c_1x_3$ becomes $c_3x_1 = kc_1c_3$, whence $x_1 = kc_1$. Likewise $x_2 = kc_2$. This completes the proof. If we had assumed $c_2 \neq 0$ or $c_1 \neq 0$ instead of $c_3 \neq 0$, the final result would have been the same.

**Example 3:** Find all the triples satisfying

$$x_1 + 2x_2 - 3x_3 = 0,$$
$$-2x_1 + 5x_2 + 4x_3 = 0.$$

We compute

$$c_1 = \begin{vmatrix} 2 & -3 \\ 5 & 4 \end{vmatrix} = 8 + 15 = 23, \qquad c_2 = \begin{vmatrix} -3 & 1 \\ 4 & -2 \end{vmatrix} = 6 - 4 = 2,$$

$$c_3 = \begin{vmatrix} 1 & 2 \\ -2 & 5 \end{vmatrix} = 5 + 4 = 9.$$

Hence the solutions are all of the multiples of (23, 2, 9).

## EXERCISES

**1.** Calculate the value of the determinant in each case.

(a) $\begin{vmatrix} 5 & 3 \\ 0 & 2 \end{vmatrix}$; (b) $\begin{vmatrix} 2 & 2 \\ 7 & 6 \end{vmatrix}$; (c) $\begin{vmatrix} 2 & 5 \\ -2 & 3 \end{vmatrix}$.

**2.** Solve by Cramer's rule:

(a) $\begin{aligned} 11x_1 - 5x_2 &= 6, \\ 3x_1 - 8x_2 &= -5. \end{aligned}$

(b) $\begin{aligned} x_1 - x_2 &= -3, \\ x_1 - 2x_2 &= 8. \end{aligned}$

(c) $\begin{aligned} 2x + 5y &= 4, \\ 3x - 4y &= -17. \end{aligned}$

(d) $\begin{aligned} 2x + 3y &= 6, \\ 4x - y &= 4. \end{aligned}$

**3.** In each case the value of the determinant is 0. Express one row as a multiple of the other row, and one column as a multiple of the other column.

(a) $\begin{vmatrix} 2 & -1 \\ 6 & -3 \end{vmatrix}$; (c) $\begin{vmatrix} 3 & 4 \\ 0 & 0 \end{vmatrix}$;

(b) $\begin{vmatrix} 4 & 6 \\ 6 & 9 \end{vmatrix}$; (d) $\begin{vmatrix} 2 & 0 \\ -3 & 0 \end{vmatrix}$.

**4.** Find all the triples satisfying the two equations in each case.

(a) $\begin{aligned} 3x_1 - 4x_2 - x_3 &= 0, \\ 5x_1 + 3x_2 + 2x_3 &= 0. \end{aligned}$

(b) $\begin{aligned} 5x_1 + x_2 - 14x_3 &= 0, \\ 7x_1 - 2x_2 + 25x_3 &= 0. \end{aligned}$

(c) $\begin{aligned} 2x - 3y + 2z &= 0, \\ 3x - y - z &= 0. \end{aligned}$

(d) $\begin{aligned} 2x + y \qquad &= 0, \\ 0 + 3y - 4z &= 0. \end{aligned}$

(e) $\begin{aligned} x + \qquad 2z &= 0, \\ 3x + y - 7z &= 0. \end{aligned}$

(f) $\begin{aligned} 2x + 3y + 2z &= 0, \\ x + 3y + z &= 0. \end{aligned}$

**5.** In the notation of (10) show that $F(x, y, u, v) = -F(u, v, x, y)$ and $F(v, u, y, x) = F(u, v, x, y)$. How are these results stated in terms of exchanges of two rows or exchanges of two columns?

6. In one of the following cases the system has no solution at all, whereas in the other case the system has many solutions. Observe the appearance of equations (11) and (12) in these cases, and describe the difference between the two cases in terms of what you observe.

(a) $\begin{aligned} 4x_1 - 2x_2 &= 8, \\ 6x_1 - 3x_2 &= 12. \end{aligned}$ (b) $\begin{aligned} 5x_1 - 15x_2 &= 3, \\ -2x_1 + 6x_2 &= -\tfrac{4}{3}. \end{aligned}$

7. In each case supply a missing number on the right side in one of the equations in such a way that the resulting pair of equations will have a solution.

(a) $\begin{aligned} 9x - 3y &= 6, \\ -12x + 4y &= \ . \end{aligned}$ (b) $\begin{aligned} 2x + 8y &= \ , \\ 3x + 12y &= 9. \end{aligned}$

8. Show that if $D = 0$ in (4), then equations (1) cannot be satisfied by a pair $(x_1, x_2)$ unless at least one of the equations is a multiple of the other. Discuss the geometric meaning of $D \neq 0$ and $D = 0$ in terms of straight lines, using $(x_1, x_2)$ instead of $(x, y)$ as coordinates. When $D = 0$, what is the *geometric* distinction between the case when the system (1) has solutions and when it does not? When it *does* have solutions, how can you describe *geometrically* the locus of all points $(x_1, x_2)$ which satisfy (1)?

9. Suppose, in a two-row determinant, that one row is a multiple of the other. Prove that one of the columns is a multiple of the other (perhaps the zero multiple).

## 17-2 Determinants of Order Three

Determinants of order three arise logically from consideration of the system of three equations

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3. \end{aligned}\right\} \quad (1)$$

by processes which are generalizations of the processes discussed in § 17-1. Our motivation is the desire to find something which corresponds to (11) and (12) in § 17-1. The plan which we follow goes like this: We eliminate $x_3$ between each pair of equations in (1), getting three equations in $x_1$ and $x_2$. Then we combine these equations in such a way as to eliminate $x_2$. The result will be what we want. The equations which we get will look a bit intricate at first. But in the process we shall be getting fundamental results which will enable us to introduce the concept of a determinant of third order.

First we multiply the second and third equations in (1) by $a_{33}$ and $-a_{23}$, respectively, and add. We symbolize this by $(2\text{nd})(a_{33}) + (3\text{rd})(-a_{23})$. Then, in like fashion, we perform $(3\text{rd})(-a_{13}) + (1\text{st})(a_{33})$ and $(1\text{st})(a_{23}) +$

(2nd)$(-a_{13})$. The resulting equations are

$$\begin{aligned}(a_{21}a_{33} - a_{31}a_{23})x_1 + (a_{22}a_{33} - a_{32}a_{23})x_2 &= b_2a_{33} - b_3a_{23},\\ (a_{11}a_{33} - a_{31}a_{13})x_1 + (a_{12}a_{33} - a_{32}a_{13})x_2 &= b_1a_{33} - b_3a_{13},\\ (a_{11}a_{23} - a_{21}a_{13})x_1 + (a_{12}a_{23} - a_{22}a_{13})x_2 &= b_1a_{23} - b_2a_{13}.\end{aligned} \tag{2}$$

Now, if we multiply these equations by $-a_{12}$, $a_{22}$, and $-a_{32}$, respectively, and add, it turns out that $x_2$ is eliminated. By using determinants of second order, the result can be written in the form

$$\left\{-a_{12}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} + a_{22}\begin{vmatrix}a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix} - a_{32}\begin{vmatrix}a_{11} & a_{13}\\ a_{21} & a_{23}\end{vmatrix}\right\} x_1$$

$$= -a_{12}\begin{vmatrix}b_2 & a_{23}\\ b_3 & a_{33}\end{vmatrix} + a_{22}\begin{vmatrix}b_1 & a_{13}\\ b_2 & a_{33}\end{vmatrix} - a_{32}\begin{vmatrix}b_1 & a_{13}\\ b_2 & a_{23}\end{vmatrix}. \tag{3}$$

This equation corresponds to (11) in § 17-1, but it needs to be expressed in a more compact and symmetrical notation. To progress toward this end let us study the coefficient of $x_1$ in (3). If we write out the actual value of each of the second order determinants, we get the following expression:

$$D = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$
$$- a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}. \tag{4}$$

There are six terms, each a product of three $a$'s. We have arranged these products so that the second indices always form 1, 2, 3, in that order. Note that the first indices always form 1, 2, 3, or some rearrangement of this set of three digits. There are, in fact, six products, corresponding to the $3! = 6$ permutations of the triple (1, 2, 3). Another significant fact is correlated with the minus signs in (4): the minus signs are on those products in which the rearrangement of 1, 2, 3 requires an *odd* number of interchanges of pairs to restore the triple to its natural order. For instance, (3, 2, 1) is restored to (1, 2, 3) by exchanging 3 and 1 (one exchange), whereas to restore (2, 3, 1) to (1, 2, 3) we must first go to (3, 2, 1) and then to (1, 2, 3), or follow some other scheme which also involves two exchanges. In the first case, $a_{31}a_{22}a_{13}$ is prefixed by a minus sign, whereas for the second case $a_{21}a_{32}a_{13}$ is not.

The number $D$ in (4) is definitely determined as a function of the $3^2$ quantities $a_{11}, a_{12}, \cdots, a_{33}$. This function is called a *determinant of order three;* the nine $a$'s are called *entries.* The standard functional notation for the determinant involves writing the entries in a square array:

$$D = \begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}. \tag{5}$$

The actual value of $D$ is defined by (4). In practice, any particular value of the determinant function is called a determinant.

With this new determinant notation we can write (3) in the form

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} x_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}. \tag{6}$$

The symmetry of the situation now strongly suggests that there are similar equations involving $x_2$ and $x_3$. There are. The equation involving $x_2$ differs from (6) merely by putting $x_2$ in place of $x_1$ and letting the column of $b$'s displace the second instead of the first column of the determinant $D$.

In order to be able to use determinants readily it is necessary for us to develop some rules which are easier to remember than formula (4). One such rule is discernible if we go back and inspect once more the coefficient of $x_1$ in (3). We also look at the display in (5). Now observe the following: if in (5) we cross out the row and column in which $a_{12}$ is located, just four entries remain, and the second-order determinant with these four entries in their natural positions is

$$\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}. \tag{7}$$

Observe also that we have $-a_{12}$ times this determinant as part of the coefficient of $x_1$ in (3). The determinant in (7) is called the *minor* of $a_{12}$ in $D$. Each entry in $D$ has a minor, which is the second-order determinant obtained by crossing out the row and column of $D$ in which that entry is located. It is convenient to have a notation for minors. We shall denote the minor of $a_{12}$ by $A_{12}$, the minor of $a_{22}$ by $A_{22}$, and so on. We now observe that the coefficient of $x_1$ in (3) [which is the same as the $D$ in (4)] can be written

$$D = -a_{12}A_{12} + a_{22}A_{22} - a_{32}A_{32}. \tag{8}$$

This formula for $D$ is called the evaluation of (or sometimes the *expansion* of) $D$ by minors of the second column.

**Example 1:** When formula (8) is applied to the calculation of

$$\begin{vmatrix} -9 & 3 & -7 \\ 6 & -4 & 4 \\ 4 & -3 & 5 \end{vmatrix},$$

we obtain

$$-3\begin{vmatrix} 6 & 4 \\ 4 & 5 \end{vmatrix} - 4\begin{vmatrix} -9 & -7 \\ 4 & 5 \end{vmatrix} + 3\begin{vmatrix} -9 & -7 \\ 6 & 4 \end{vmatrix}$$

$$= -3(30 - 16) - 4(-45 + 28) + 3(-36 + 42) = 44.$$

The rule expressed in (8) does relieve us of the burden of trying to remember the formidable-looking formula (4). But (8) also raises some natural questions. Why is the second column especially important—or is it? Why do we prefix $a_{12}$ and $a_{32}$, but not $a_{22}$, by minus signs in (8)? To see what we can learn about answers to these questions, let us go back to examine the way in which we derived formula (3). We began by eliminating $x_3$ from the system (1), our first step being the obtaining of equations (2). Thereafter we eliminated $x_2$. Suppose we had done things in reverse order as regards $x_2$ and $x_3$. By eliminating first $x_2$ and then $x_3$, in a pattern analogous to that of our original procedure, we would have come out with the following equation:

$$\left\{ a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - a_{23}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \right\} x_1$$

$$= a_{13}\begin{vmatrix} b_2 & a_{22} \\ b_3 & a_{32} \end{vmatrix} - a_{23}\begin{vmatrix} b_1 & a_{12} \\ b_3 & a_{33} \end{vmatrix} + a_{33}\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}.$$

This looks different from (3), but if we calculate the value of each side, we find that it is in fact the same as (3). The coefficient of $x_1$ here is the same $D$ as in (4), but now it is expressed as

$$D = a_{13}A_{13} - a_{23}A_{23} + a_{33}A_{33}. \tag{9}$$

As in the case of (8), we have here an evaluation of $D$, this time by minors of the *third* column. Now, however, the only entry which is prefixed by a minus sign is $a_{23}$.

If we were to solve for $x_2$ by eliminating first $x_3$ and then $x_1$ by the same pattern as was used before, we would obtain still another evaluation of $D$, by minors of the first column:

$$D = a_{11}A_{11} - a_{21}A_{21} + a_{31}A_{31}.$$

The explanation of the minus signs in all cases is this: In an evaluation of $D$ by minors of a particular column, an entry from that column is to be prefixed by a minus sign if and only if the sum of its indices is odd, that is, if and only if the sum of its row number and its column number is odd.

**Example 2:** If the determinant in Example 1 is evaluated by minors of the first column, the calculations are as follows:

$$-9\begin{vmatrix} -4 & 4 \\ -3 & 5 \end{vmatrix} - 6\begin{vmatrix} 3 & -7 \\ -3 & 5 \end{vmatrix} + 4\begin{vmatrix} 3 & -7 \\ -4 & 4 \end{vmatrix}$$

$$= -9(-20 + 12) - 6(15 - 21) + 4(12 - 28) = 44.$$

## EXERCISES

1. Calculate the value of the third-order determinant at least twice in each case, using minors of one column and then of another.

(a) $\begin{vmatrix} 2 & 3 & -5 \\ 1 & -2 & 1 \\ 3 & 1 & 1 \end{vmatrix}$.

(b) $\begin{vmatrix} 2 & -1 & 4 \\ 7 & 5 & -2 \\ -3 & 2 & 4 \end{vmatrix}$.

(c) $\begin{vmatrix} 5 & -2 & -3 \\ 2 & 4 & -1 \\ 7 & 2 & -4 \end{vmatrix}$.

(d) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 0 & 1 & 2 \end{vmatrix}$.

(e) $\begin{vmatrix} 1 & 3 & -2 \\ 1 & 2 & -2 \\ -1 & -1 & 2 \end{vmatrix}$.

(f) $\begin{vmatrix} 2 & 3 & 1 \\ 4 & 1 & 0 \\ -1 & 2 & -2 \end{vmatrix}$.

(g) $\begin{vmatrix} -1 & 1 & 3 \\ 1 & 0 & 1 \\ 1 & -2 & 2 \end{vmatrix}$.

(h) $\begin{vmatrix} 4 & 5 & -2 \\ 7 & -4 & 5 \\ 7 & 2 & 1 \end{vmatrix}$.

**2.** Work out in detail the derivation of the equation analogous to (6) with $x_2$ in place of $x_1$, by starting with (1) and eliminating first $x_3$ and then $x_1$, in a pattern similar to that employed in arriving at equation (3).

## 17-3 Further Discussion of Third-Order Determinants

A number of significant observations about determinants can be made on the basis of what was said in § 17-2.

THEOREM 17-B. *Consider two third-order determinants whose schematic arrays are related to each other as follows: each row of one is the same as the corresponding column of the other. Then these two determinants are equal in value.*

*Proof.* Consider the determinant (5) and its value (4) in § 17-2. If we construct another determinant whose columns are the rows of (5), and denote its entries by $b_{11}$, $b_{12}$, and so on, then $b_{ij} = a_{ji}$. That is, $b_{11} = a_{11}$, $b_{12} = a_{21}$, and so on. But now, if we examine (4) carefully, we see that if we were to write a similar expression in which each $a_{ij}$ has its indices exchanged, the total expression would be just the same as before. The second and the third product in (4) would merely be exchanged, and the last three products would merely have their factors rearranged. This proves the theorem.

It follows from this theorem that a determinant can be evaluated by minors of any selected row, as well as by minors of a given column. The rule of signs for the evaluation by minors of a row is the same as in the evaluation by minors of a column.

THEOREM 17-C. *A determinant has the value* 0 *if some row in it is a multiple of another row, or if some column is a multiple of another column.*

*Proof.* It will be sufficient to consider the case of rows. The case of columns is then taken care of by applying Theorem 17-B. If one row is a multiple, say by the factor $k$, of another row, let us evaluate the determinant by minors of the remaining row. For instance, if the second row is $k$ times the third row, then we evaluate by minors of the first row. Then each of the minors is a second-order determinant which is 0, because one of its rows is a multiple of another row. (This fact about second-order determinants was mentioned in § 17-1). But if the minors are zero, so is the value of the whole determinant.

Next, it is fruitful to think of the columns of a determinant as entities, and to consider some simple facts about how the value of the determinant depends upon one of its columns. Now a column is a triple of numbers. We shall have occasion to speak of *linear combinations* of triples. By the *sum* of two triples $(x, y, z)$ and $(u, v, w)$ we mean the triple $(x + u, y + v, z + w)$. By $k$ times the triple $(x, y, z)$ we mean the triple $(kx, ky, kz)$. By a linear combination of two or more triples we mean a sum of multiples of the several triples.

THEOREM 17-D. *Consider a determinant, one of whose columns is formed as a linear combination of two triples. Then the value of the determinant is this same linear combination of the values of the two determinants which result by replacing the original column, first by one of the triples, and then by the other.*

Before giving the proof let us illustrate the meaning of the theorem by an example in which the second column is the one considered. The theorem asserts that

$$\begin{vmatrix} a_{11} & ax + bu & a_{13} \\ a_{21} & ay + bv & a_{23} \\ a_{31} & az + bw & a_{33} \end{vmatrix} = a\begin{vmatrix} a_{11} & x & a_{13} \\ a_{21} & y & a_{23} \\ a_{31} & z & a_{33} \end{vmatrix} + b\begin{vmatrix} a_{11} & u & a_{13} \\ a_{21} & v & a_{23} \\ a_{31} & w & a_{33} \end{vmatrix}. \tag{1}$$

The proof comes directly from an examination of the formula (4) in § 17-2. Each of the products in this formula is seen to involve just one entry from the second column, and each product therefore behaves in the proper manner. For example, if in the product $a_{21}a_{32}a_{13}$ we replace $a_{32}$ by $az + bw$, the result is $a(a_{21}z\, a_{13}) + b(a_{21}w\, a_{13})$. When the corresponding thing is done for each product and the results are examined, we see that we have a proof of (1). The proof for the case of some other column is made in exactly the same way.

**Example 1:**

$$\begin{vmatrix} 1 & 3+6 & 1 \\ 2 & 3+4 & -1 \\ 3 & 3+2 & 1 \end{vmatrix} = 3\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{vmatrix} + 2\begin{vmatrix} 1 & 3 & 1 \\ 2 & 2 & -1 \\ 3 & 1 & 1 \end{vmatrix}.$$

The foregoing theorem is used in proving our next result.

THEOREM 17-E. *In a given determinant let us select two columns, as for example, the first and third. Then let us form a new determinant as follows: it shall have the same second and third columns as the original ones, but its first column shall be the sum of the original first column and any chosen multiple of the original third column. Then this new determinant is equal in value to the original one. The result is general; that is, it applies to any two columns.*

*Proof.* For definiteness we assume that $k$ times the third column is added to the first column to form the new first column. Let the value of the original determinant be $D$, that of the new one $D_1$. Then, by Theorem 17-D, we have

$$D_1 = D + kD_0,$$

where $D_0$ is a determinant whose first and third columns are the same as the third column of $D$. But $D_0 = 0$, by Theorem 17-C. Therefore $D_1 = D$, as asserted.

**Example 2:**

$$\begin{vmatrix} 2 & 1 & 7 \\ -3 & 5 & -2 \\ 4 & 3 & 0 \end{vmatrix} = \begin{vmatrix} 2-7 & 1 & 7 \\ -3+2 & 5 & -2 \\ 4-0 & 3 & 0 \end{vmatrix}.$$

The process described in Theorem 17-E can be used repeatedly. Its worth is that we may be able to simplify the calculation of the determinant by getting a column which has several zero entries. This shortens the evaluation by minors.

**Example 3:** Here we shall add 3 times the third column to the first column:

$$\begin{vmatrix} 3 & 2 & -1 \\ 4 & 3 & 2 \\ 6 & -5 & -2 \end{vmatrix} = \begin{vmatrix} 0 & 2 & -1 \\ 10 & 3 & 2 \\ 0 & -5 & -2 \end{vmatrix} = -10 \begin{vmatrix} 2 & -1 \\ -5 & -2 \end{vmatrix}$$

$$= -10(-4-5) = 90.$$

If all the entries in a single column are 0, the value of the determinant is 0. Hence we see, by repeated application of Theorem 17-E, that the value of a determinant is 0 if we can form a linear combination of its columns, using a nonzero multiple of at least one column, so as to obtain a column of zeros. When this is the case we say that the columns of the determinant are linearly dependent, or that one of them is a linear combination of the others.

**Example 4:** The determinant

$$\begin{vmatrix} 2 & 4 & 10 \\ -3 & 3 & 12 \\ 5 & 4 & 7 \end{vmatrix}$$

is equal to 0. The linear dependence is

first column = 3(second column) − (third column).

There is a linear dependence of rows also. It is

9(first row) − 4(second row) − 6(third row) = row of zeros.

Ways of discovering this sort of linear dependence, when it exists, are considered in the last part of the next section.

## EXERCISES

1. If two columns of a determinant are interchanged, the value of the determinant is replaced by the negative of the original value. This can be proved as follows, using Theorems 17-C and 17-D. Suppose that the second and third columns are to be exchanged. Start from

$$0 = \begin{vmatrix} a_{11} & a_{12} + a_{13} & a_{12} + a_{13} \\ a_{21} & a_{22} + a_{23} & a_{22} + a_{23} \\ a_{31} & a_{32} + a_{33} & a_{32} + a_{33} \end{vmatrix},$$

a relation which is true by Theorem 17-C. Now use Theorem 17-D several times, and also Theorem 17-C, until the result

$$0 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{vmatrix}$$

is reached. Write out every step in detail.

2. Prove that

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{13} & -a_{12} \\ a_{21} & a_{23} & -a_{22} \\ a_{31} & a_{33} & -a_{32} \end{vmatrix}$$

by repeated use of Theorem 17-E. How does the result of Exercise 1 then follow?

3. Show that

$$\begin{vmatrix} 1 & a+x & a-x \\ 2 & b+y & b-y \\ 3 & c+z & c-z \end{vmatrix} = -2\begin{vmatrix} 1 & a & x \\ 2 & b & y \\ 3 & c & z \end{vmatrix}.$$

4. Explain, *on the basis of theorems in this section,* why each of the following determinants has the value 0.

(a) $\begin{vmatrix} 1 & 2 & 5 \\ 2 & 4 & -1 \\ -3 & -6 & 2 \end{vmatrix}.$ (d) $\begin{vmatrix} 1 & 2 & -1 \\ 3 & -1 & 4 \\ 4 & 1 & 3 \end{vmatrix}.$

(b) $\begin{vmatrix} 1 & 0 & -1 \\ 1 & 2 & -1 \\ -1 & 3 & 1 \end{vmatrix}$.

(e) $\begin{vmatrix} 1 & 3 & 2 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix}$.

(c) $\begin{vmatrix} 2 & 4 & 1 \\ 3 & 5 & 2 \\ 6 & 12 & 3 \end{vmatrix}$.

(f) $\begin{vmatrix} -6 & 15 & 3 \\ 5 & -4 & 2 \\ 3 & 1 & 3 \end{vmatrix}$.

**5.** Explain on the basis of theorems in this section, why each of the following pairs of determinants are equal.

(a) $\begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{vmatrix} = \begin{vmatrix} 3 & 6 & 9 \\ 1 & 4 & 7 \\ 2 & 5 & 8 \end{vmatrix}$.

(b) $\begin{vmatrix} 1 & 3 & 2 \\ 1 & 3 & 4 \\ 2 & 4 & 6 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 \\ 3 & -3 & 4 \\ 4 & -2 & 6 \end{vmatrix}$.

(c) $\begin{vmatrix} 2 & 4 & 1 \\ 3 & 5 & 2 \\ 6 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 2 & 4 & -3 \\ 3 & 5 & -4 \\ 6 & 1 & -6 \end{vmatrix}$.

**6.** (a) See if you can discover the linear dependence of the columns which insures that

$$\begin{vmatrix} 4 & 5 & -2 \\ 7 & -4 & 5 \\ 7 & 2 & 1 \end{vmatrix} = 0.$$

(b) Can you discover also the linear dependence of the rows?

**7.** Calculate the value of each determinant by methods analogous to that of Example 3.

(a) $\begin{vmatrix} 10 & 15 & 20 \\ 12 & 12 & 32 \\ 2 & 3 & 12 \end{vmatrix}$.

(c) $\begin{vmatrix} 3 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & -2 & 1 \end{vmatrix}$.

(b) $\begin{vmatrix} 3 & 2 & 1 \\ 3 & -3 & 2 \\ 10 & 1 & 7 \end{vmatrix}$.

(d) $\begin{vmatrix} 2 & 4 & -1 \\ 3 & 1 & 2 \\ 1 & 0 & -2 \end{vmatrix}$.

**8.** (a) Show that

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ a-b & b-c & c \\ a^2-b^2 & b^2-c^2 & c^2 \end{vmatrix},$$

and hence that the value of the determinant is $(a - b)(b - c)(c - a)$.
(b) Show that

$$\begin{vmatrix} 1 & a & a^3 \\ 1 & b & b^3 \\ 1 & c & c^3 \end{vmatrix} = (a - b)(b - c)(c - a)(a + b + c).$$

## 17-4 The Solution of Linear Systems

We go back now to the discussion of the system of three equations in three unknowns as presented in (1), § 17-2. Our discussion has to do, not so much with the practical problem of finding solutions of a system of this kind in particular numerical cases, as with the general question of whether there are any solutions at all of the system, and if there *are* solutions, whether there is uniqueness of solution.

To illustrate the possibility of there being no solution at all, consider the equations

$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 1, \\ 4x_1 + 6x_2 - 2x_3 &= 4, \\ x_1 - x_2 + x_3 &= 2. \end{aligned}$$

It can be seen right away that there is no triple $(x_1, x_2, x_3)$ which satisfies all three equations. If there were, the first two equations would give contradictory results. For, on multiplying the first equation by 2, we see that $4x_1 + 6x_2 - 2x_3 = 2$, whereas the second equation demands that $4x_1 + 6x_2 - 2x_3 = 4$, not 2.

The determinant $D$ in (5) of § 17-2, formed with the coefficients of the linear system as entries in the manner shown, is called *the determinant of the linear system* (1) of § 17-2.

**Example 1:** The determinant of the system

$$\begin{aligned} 2x - 3y + z &= 4, \\ x + y - z &= 2, \\ 4x - y + 3z &= 1, \end{aligned}$$

is

$$\begin{vmatrix} 2 & -3 & 1 \\ 1 & 1 & -1 \\ 4 & -1 & 3 \end{vmatrix}.$$

By a solution of (1) in § 17-2 we mean a triple of numbers $(x_1, x_2, x_3)$ which satisfies all three equations.

THEOREM 17-F. *The system* (1) *in* § 17-2 *has a unique solution if* $D \neq 0$.

*This solution is given by*

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{D} \tag{1}$$

*and two similar equations for $x_2$, $x_3$, as described in connection with* (6) *in* § 17-2.

*Proof.* If $D \neq 0$ and if there is a solution, the formula here given for $x_1$ is an immediate consequence of (6) in § 17-2. The situation for $x_2$ and $x_3$ is similar. Hence, when $D \neq 0$ the solution, if it exists, is unique. To prove that the solution really does exist, we define $x_1$ by the formula in the theorem, and $x_2$, $x_3$ likewise; then we verify by actual substitution that these values of $x_1$, $x_2$, $x_3$ do satisfy the linear system. We give the details merely for the first equation, since this illustrates the way in which the verification is made.

What we wish to verify is that

$$\frac{a_{11}}{D}\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix} + \frac{a_{12}}{D}\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix} + \frac{a_{13}}{D}\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} = b_1.$$

We shall evaluate each determinant by minors of the column in which the $b$'s are located. The left side of the foregoing equation then becomes

$$\frac{a_{11}}{D}[b_1A_{11} - b_2A_{21} + b_3A_{31}] + \frac{a_{12}}{D}[-b_1A_{12} + b_2A_{22} - b_3A_{32}] + \frac{a_{13}}{D}[b_1A_{13} - b_2A_{23} + b_3A_{33}].$$

Now, the coefficient of $b_1$ in all of this is

$$\frac{1}{D}[a_{11}A_{11} - a_{12}A_{12} + a_{13}A_{13}] = \frac{D}{D} = 1,$$

for the bracketed expression here is exactly the evaluation of $D$ by minors of the first row. What remains, then, is to show that the coefficients of $b_2$ and $b_3$ are 0. We examine the coefficient of $b_2$; the case for $b_3$ is similar. The coefficient of $b_2$ is

$$\frac{1}{D}[-a_{11}A_{21} + a_{12}A_{22} - a_{13}A_{23}].$$

The $A$'s here are the minors of the entries in the second row of $D$. Hence the expression in brackets is the evaluation, by minors of the second row, of the determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

This determinant is 0 by Theorem 17-C, for the first and second rows are the same. This completes the proof of Theorem 17-F.

The formulas which express $x_1$, $x_2$, $x_3$ as quotients of determinants (as with $x_1$ in the statement of Theorem 17-F) are jointly known as *Cramer's rule* (after an 18th century Swiss mathematician). These formulas are more of theoretical interest than of practical value for computation, because the solution can usually be found with less computational labor than is involved in calculating all the determinants which appear in Cramer's rule.

*Homogeneous Linear Systems*

Consider the equations

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= 0, \end{aligned}\right\} \quad (2)$$

in which all three right members are 0. This is called a *homogeneous system* (see the remarks made in connection with (13) in § 17-1). If the determinant $D$ of the system (2) is not 0, the unique solution of the system is $x_1 = x_2 = x_3 = 0$, by Cramer's rule, because

$$x_1 = \frac{\begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix}}{D} = 0,$$

with similar results for $x_2$ and $x_3$. But if $D = 0$, there are solutions of the system (2) in which $x_1$, $x_2$, $x_3$ are not all 0. We shall prove this in a moment. But we observe that this implies that there is no uniqueness about the solution of (2). For, if $(x_1, x_2, x_3)$ is a solution, so is $(2x_1, 2x_2, 2x_3)$, and this second solution is different from the first one if at least one of the $x$'s is not 0.

How shall we find a solution [other than (0, 0, 0)] of (2) when $D = 0$? There are two cases to consider. The first case is that in which there is at least one entry in $D$ whose corresponding minor is not 0. This minor then involves two rows of the determinant $D$, and we consider the equations corresponding to these two rows. To these two equations we apply the method of solution explained in the proof of Theorem 17-A. For example, suppose the two rows in question are the first and second. According to Theorem 17-A, the only solutions of the first two equations are multiples of $(x_1, x_2, x_3)$, where

$$x_1 = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = A_{31}, \qquad x_2 = \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix} = -A_{32}, \qquad x_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = A_{33}.$$

Now, if we expand $D$ by minors of the third row, we get

$$0 = D = a_{31}A_{31} - a_{32}A_{32} + a_{33}A_{33}.$$

This shows that our values for $x_1$, $x_2$, $x_3$ also satisfy the third equation in (2). The situation could be handled similarly, starting with a different pair of equations, if the corresponding minors were not all 0.

In the alternative case all the minors of elements of $D$ are 0. This implies that each pair of rows of the determinant are linearly dependent, and hence that there is one row such that the other two are multiples of it. In this case we have only to solve one of the equations (2); the other two will then be satisfied automatically. The solution of a single equation in three unknowns is naturally not unique. In general we can assign arbitrary values to two of the letters $x_1$, $x_2$, $x_3$ and then solve for the third.

**Example 2:** Find all solutions of the system

$$\begin{aligned} 2x_1 - \phantom{1}3x_2 + 5x_3 &= 0, \\ 4x_1 + \phantom{1}3x_2 + 4x_3 &= 0, \\ 10x_1 + 12x_2 + 7x_3 &= 0. \end{aligned}$$

The determinant of this system is 0 (it is obtained from the determinant of Example 4, § 17-3, by exchanging rows and columns). We solve the first and second equations by the method of Theorem 17-A. The result is

$$x_1 = \begin{vmatrix} -3 & 5 \\ 3 & 4 \end{vmatrix} = -27, \qquad x_2 = \begin{vmatrix} 5 & 2 \\ 4 & 4 \end{vmatrix} = 12, \qquad x_3 = \begin{vmatrix} 2 & -3 \\ 4 & 3 \end{vmatrix} = 18.$$

Hence $(-27, 12, 18)$ is a solution. Hence so is $-\frac{1}{3}$ of this, or $(9, -4, -6)$. The third equation is automatically satisfied. All other solutions are multiples of the basic $(9, -4, -6)$.

Observe that this result implies the following linear dependence of columns:

$$9(\text{1st column}) - 4(\text{2nd column}) - 6(\text{3rd column}) = \text{a column of zeros.}$$

**Example 3:** Consider the system

$$\begin{aligned} 4x_1 - 2x_2 - 6x_3 &= 0, \\ 2x_1 - \phantom{2}x_2 - 3x_3 &= 0, \\ 6x_1 - 3x_2 - 9x_3 &= 0. \end{aligned}$$

In this case the first and third equations are multiples of the second, by factors 2 and 3, respectively. Here $D$ and the minors of all its entries are zero. For a solution, we may assign $x_1$ and $x_3$ arbitrarily, and calculate $x_2$ by $x_2 = 2x_1 - 3x_3$.

For a geometric interpretation of Examples 2 and 3 we must wait for the study of lines and planes in Chapter XVIII.

## EXERCISES

**1.** Write the solution of each system by Cramer's rule, using quotients of determinants. Calculate the determinants and so obtain the solution.

(a) $\begin{cases} 3x - 8y + 6z = 1, \\ 2x + 4y - 3z = 3, \\ 8x - 2y - 9z = 4. \end{cases}$ (c) $\begin{cases} x + y + z = 9, \\ x + 2y + 3z = 9, \\ x + 3y + 6z = 3. \end{cases}$

(b) $\begin{cases} x_1 + 4x_2 + 5x_3 = 9, \\ 2x_1 + 3x_3 = 13, \\ 3x_1 + 9x_3 = 33. \end{cases}$ (d) $\begin{cases} 3x - 2z = 10, \\ -2x + 3y = 12, \\ -2y + 3z = -23. \end{cases}$

**2.** Find all solutions of each system.

(a) $\begin{cases} x_1 + 2x_2 - x_3 = 0, \\ 3x_1 - 4x_2 + 2x_3 = 0, \\ x_1 + 12x_2 - 6x_3 = 0. \end{cases}$ (c) $\begin{cases} 2x + y + 2z = 0, \\ x + 5y - 5z = 0, \\ -x - 2y + z = 0. \end{cases}$

(b) $\begin{cases} 3x - 3y - z = 0, \\ x - y - 4z = 0, \\ 2x - 2y + z = 0. \end{cases}$ (d) $\begin{cases} 3x + 4y + z = 0, \\ 5x - y + 2z = 0, \\ 8x + y - 2z = 0. \end{cases}$

**3.** The determinant

$$\begin{vmatrix} 3 & 2 & -2 \\ 2 & 3 & -1 \\ 8 & 7 & -5 \end{vmatrix}$$

has the value 0. (a) Find the linear dependence of its columns by solving an appropriate homogeneous linear system. (b) Find the linear dependence of the rows by solving a related homogeneous linear system.

**4.** Show that three points $(x_1, y_1)$, $(x_2, y_2)$, $(x_3, y_3)$ in the $xy$-plane lie on a single straight line if and only if

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

*Suggestion:* For the points to lie on a line it is necessary and sufficient that there exist numbers $A$, $B$, $C$, not all 0, such that all three points satisfy the equation $Ax + By + C = 0$.

5. Use the result in Exercise 4 to show that the equation of the straight line through the two distinct points $(x_1, y_1)$, $(x_2, y_2)$ is

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$$

## 17-5 Determinants of Higher Order

We shall be very brief, and merely suggest what the general situation is for determinants of order $n$ if $n > 3$. The main ideas have been touched on in dealing with the case $n = 3$. Everything is arranged so that determinants of order $n$ have the same relation to systems of $n$ linear equations in $n$ unknowns as has already been developed for $n = 2$ and $n = 3$. We express the determinant schematically in the array

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Its value can be worked out by using minors of any selected row or column with the same rule of signs as in the case $n = 3$. The minors are determinants of order $n - 1$. Thus, for the evaluation of a determinant of order 4, we shall have a sum of four multiples of determinants of order 3. What corresponds to formula (4) in § 17-2 is an algebraic sum of $n!$ products, each product involving $n$ different entries, one from each row and one from each column. The rule of signs for these products is expressed in a manner which we shall not discuss. Suffice it to say here that it depends upon a study of permutations of the natural order of the numbers $1, 2, \cdots, n$, and is a natural extension of the rule explained at the end of § 17-2.

# CHAPTER XVIII

# ANALYTIC GEOMETRY OF THREE DIMENSIONS

## 18-1 Fundamental Notions

In § 6-6 we described the way in which a rectangular coordinate system is introduced for the purpose of discussing the location of points in space of three dimensions. Each point $P$ is identified by its coordinates $(x, y, z)$ in this coordinate system. The correspondence between $P$ and its coordinates is a one-to-one correspondence between the totality of points in space and the totality of ordered triples of real numbers. All aspects of the geometry of space can be studied through the medium of the coordinate system, but it is often better for the sake of directness, clarity, and intuitive perception, to develop familiarity with geometric objects and geometric relationships which can be thought of without the intervention of coordinates.

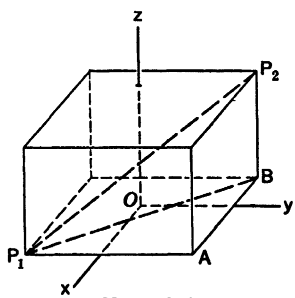


Fig. 18-1

*The Distance Formula*

If $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ are any two points, the square of the distance between them is

$$d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2. \tag{1}$$

This can be worked out by using the theorem of Pythagoras twice. In Fig. 18-1, $P_1AB$ is a right triangle with right angle at $A$, and $P_1BP_2$ is a right triangle with right angle at $B$. This is because the box has been constructed with $P_1$ and $P_2$ at diagonally opposite corners and with each face of the box parallel to a coordinate plane. Then

$$\overline{P_1P_2}^2 = \overline{P_1B}^2 + \overline{BP_2}^2, \qquad \overline{P_1B}^2 = \overline{P_1A}^2 + \overline{AB}^2.$$

Since $\overline{AB} = |x_2 - x_1|, \quad \overline{P_1A} = |y_2 - y_1|, \quad \overline{BP_2} = |z_2 - z_1|,$

the combination of these results yields the formula (1).

*Spheres*

It is immediate from (1) that a sphere (i.e., the surface of the sphere) with center $(a, b, c)$ and radius $r$ is characterized by the equation

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = r^2. \qquad (2)$$

Many problems with spheres are similar to corresponding problems with circles, so far as the algebra of the problems is concerned. In particular, the technique of completing the square is useful in locating the center of a sphere.

**Example 1:** Consider the locus of all points $P(x, y, z)$ such that the distance from $P$ to the origin is one third of the distance from $P$ to $(8, 0, 0)$. Show that the locus is a sphere, and find its center and radius.

The condition on the locus is that

$$(x^2 + y^2 + z^2)^{1/2} = \frac{1}{3}[(x - 8)^2 + y^2 + z^2]^{1/2}.$$

Squaring both sides and simplifying, we bring this to the form

$$8(x^2 + y^2 + z^2) + 16x = 64.$$

The completion of square technique then gives

$$x^2 + 2x + 1 + y^2 + z^2 = 8 + 1,$$

or

$$(x + 1)^2 + y^2 + z^2 = 9.$$

Hence the locus is a sphere with radius 3 and center at $(-1, 0, 0)$.

*Vectors*

We can easily extend to three dimensions the ideas about vectors as they were presented in § 13-1. A vector as a geometric object is a directed line segment from the origin to some point; we imagine it as fitted with a tip at the end. The zero vector **O**, just as before, is merely the origin itself. The addition of two vectors is defined exactly as before, and so is the process of multiplying a vector by a number. Observe that, when just two vectors are involved, they either lie along the same line or they determine a plane, the plane through the origin and the tips of the two vectors. The two vectors and their sum then lie in this plane.

Vectors of unit length in the positive directions along the axes are called the *standard unit vectors*. We denote them by **i**, **j**, **k**. Any vector can be expressed as a linear combination of these standard vectors. If **R** is a vector whose tip is $P(x, y, z)$, then

$$\mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}. \tag{3}$$

This is shown in Fig. 18-2. It is often convenient to refer to **R** as the vector $(x, y, z)$. The coordinates $x$, $y$, $z$ are called *components* of **R**.

We denote the length of **R** by $|\mathbf{R}|$.

*Parametric Representation of a Straight Line*

Let $L$ be a complete straight line anywhere in space. For many purposes it is convenient to think of a line as being determined by a point and a direction through that point. Using this idea, we can visualize any point

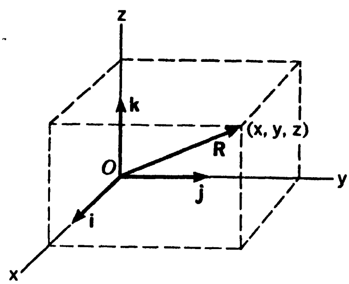



**Fig. 18-2**

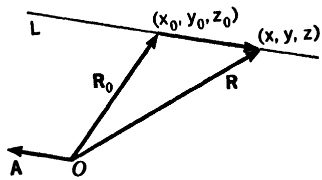



**Fig. 18-3**

on the line as being reached by the sum of two vectors, as follows: Let $P_0(x_0, y_0, z_0)$ be a point on the line and let **A** be a nonzero vector which is parallel to the line. Then, if **R** is the vector from $O$ to any point $(x, y, z)$ on the line, **R** is the sum of the vector $\mathbf{R}_0$ from $O$ to $(x_0, y_0, z_0)$ and a certain multiple of the vector **A**, so that

$$\mathbf{R} = \mathbf{R}_0 + t\mathbf{A}. \tag{4}$$

The situation is shown in Fig. 18-3; for this case $t$ would be negative. If **A** has components $(a, b, c)$, the vector formula (4) is equivalent to the three equations

$$x = x_0 + at, \qquad y = y_0 + bt, \qquad z = z_0 + ct, \tag{5}$$

which are parametric equations of the line.

Conversely, a set of equations of the form (5) in which $a$, $b$, $c$ are not all 0 can be put back into the vector form (4), and therefore represents a straight line through $(x_0, y_0, z_0)$ with direction parallel to the vector with components $(a, b, c)$.

**Example 2:** A line goes through $(2, -3, 4)$ and is parallel to the vector with components $(4, -3, 2)$. Find where the line pierces the $xz$-plane.

As parametric equations of the line we have

$$x = 2 + 4t, \qquad y = -3 - 3t, \qquad z = 4 + 2t.$$

The $xz$-plane is characterized by the equation $y = 0$. We therefore set $-3 - 3t = 0$ and solve for $t$, getting $t = -1$. When this is put back in the parametric equations, we find $x = -2$, $y = 0$, $z = 2$ for the required point.

## EXERCISES

**1.** In Fig. 18-1 let $P_1$ be $(4, -2, 2)$ and let $P_2$ be $(-1, 3, 6)$. (a) Write the equations of the six faces of the box shown in Fig. 18-1, arranging the faces in pairs perpendicular to the $x$-axis, the $y$-axis, and the $z$-axis, respectively. (b) Find the volume of the box. (c) What equations describe the line through $P_1$ and $A$? The line through $A$ and $B$? The line through $B$ and $P_2$?

**2.** A rectangular box has its faces in the planes $x = 1$, $x = 7$, $y = 3$, $y = 5$, $z = 3$, $z = 8$, respectively. (a) Sketch the box. (b) Find the coordinates of its corners. (c) Find the dimensions of the box, its volume, and the length of one of its diagonals.

**3.** (a) Find the perimeter of the triangle with vertices $(3, 1, -2)$, $(1, -4, 2)$, $(-1, 3, 3)$.
(b) Show that the triangle with vertices $(5, 4, 7)$, $(-1, 1, 9)$, $(2, 6, 1)$ is a right triangle. Find its area.

**4.** Using distances, determine which of the following sets of points are collinear.
(a) $(3, -2, 5)$, $(9, 1, -1)$, $(13, 3, -5)$.
(b) $(6, -1, -5)$, $(4, 2, -2)$, $(-2, 8, 4)$.
(c) $(6, 2, -2)$, $(3, 6, 0)$, $(0, 10, 3)$.
(d) $(3, 0, -3)$, $(7, 8, 5)$, $(10, 14, 11)$.

**5.** (a) Find a point on the $y$-axis which is equidistant from $(2, 4, -3)$ and $(-3, 5, 1)$. (b) Find a point in the plane $x = 0$ and equidistant from $(3, 0, 2)$, $(2, 3, 0)$ and $(1, 0, 0)$.

**6.** (a) How far is $(-3, 5, 1)$ from the $y$-axis? (b) How far is $(2, 4, 6)$ from the line through $(1, 6, 8)$ parallel to the $x$-axis?

**7.** Explain how to find the coordinates of the mid-point of the line segment $P_1P_2$, given the coordinates of $P_1$ and $P_2$. *Suggestion:* Drop perpendiculars from $P_1$ and $P_2$ to the $xy$-plane, obtaining points $Q_1$, $Q_2$. What are their coordinates? Explain why the mid-point of $Q_1Q_2$ has the same $x$ and $y$ coordinates as the mid-point of $P_1P_2$.

**8.** Find the equation of the sphere having the two given points as ends of a diameter.
(a) $(7, 3, 5)$, $(-1, 6, -1)$. (b) $(2, -1, 3)$, $(5, 5, 9)$.

**9.** Write the equation of the sphere of radius 5 with center on the positive $y$-axis and tangent to the plane $y = 0$.

10. Find the locus of points each of whose distance from (4, 4, 0) is twice its distance from (0, −2, 1). Identify the locus as a sphere, and find its center and radius.

11. Identify the locus by name, and if it is a sphere, tell its center and radius. In some cases there may be no locus.
   (a) $x^2 + y^2 + z^2 - 2x + 4y + 6z = 2$.
   (b) $x^2 + y^2 + z^2 - 6x + 8y + 4z + 29 = 0$.
   (c) $x^2 + y^2 + z^2 - 8x - 2y + 4z = 4$.
   (d) $x^2 + y^2 + z^2 + 4x - 6y + 12z + 61 = 0$.
   (e) $36(x^2 + y^2 + z^2) - 36x + 48y - 84z + 5 = 0$.
   (f) $9(x^2 + y^2 + z^2) + 12x + 6y + 5 = 0$.

12. Find and simplify an equation describing the locus of all points equidistant from (4, 2, −3) and (−2, 0, 7). Describe the locus in geometric language.

13. What vector must be added to **A** to give **B** if (a) $\mathbf{A} = 2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}$, $\mathbf{B} = 3\mathbf{i} - 3\mathbf{j} + 8\mathbf{k}$; (b) $\mathbf{A} = 3\mathbf{i} + 5\mathbf{j} - 6\mathbf{k}$, $\mathbf{B} = 6\mathbf{i} - \mathbf{j} + 2\mathbf{k}$?

14. Using vectors, find two points on the line through $A(1, 2, 3)$ and $B(3, 4, 2)$ which are twice as far from $A$ as from $B$. Start by making a sketch with the plane of your paper representing the plane through $O$, $A$, and $B$.

15. A line $L$ goes through (−2, 7, 4) and is parallel to the vector $6\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$. Find the point on the line nearest the origin.

16. A line passes through the points (2, 1, 2) and (8, −1, −8). Find (a) where it intersects the $xz$-plane; (b) where it intersects the plane $x = -1$; (c) the point on the line closest to the origin.

## 18-2 The Angle Between Two Vectors. The Scalar Product

Consider the two nonzero vectors

$$\mathbf{A}_1 = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}, \qquad \mathbf{A}_2 = a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}.$$

We wish to deduce a formula for the angle $\theta$ between the vectors (see Fig. 18-4). By the distance formula

$$d^2 = (a_2 - a_1)^2 + (b_2 - b_1)^2 + (c_2 - c_1)^2. \quad (1)$$

Now, the length of $\mathbf{A}_1$ is

$$|\mathbf{A}_1| = (a_1^2 + b_1^2 + c_1^2)^{1/2},$$

and there is a similar formula for the length of $\mathbf{A}_2$. Therefore, by the law of cosines,

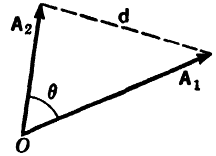


**Fig. 18-4**

$$d^2 = a_1^2 + b_1^2 + c_1^2 + a_2^2 + b_2^2 + c_2^2$$
$$- 2(a_1^2 + b_1^2 + c_1^2)^{1/2}(a_2^2 + b_2^2 + c_2^2)^{1/2} \cos\theta.$$

If, in this expression, we replace $d^2$ by its value from (1), we obtain the formula

$$\cos\theta = \frac{a_1a_2 + b_1b_2 + c_1c_2}{(a_1^2 + b_1^2 + c_1^2)^{1/2}(a_2^2 + b_2^2 + c_2^2)^{1/2}}. \tag{2}$$

This is the formula we wished to find.

**Example 1:** Find the angle $\theta$ between the vectors

$$\mathbf{A} = 3\mathbf{i} + 2\mathbf{j} + 6\mathbf{k}, \qquad \mathbf{B} = 2\mathbf{i} + 4\mathbf{j} - 4\mathbf{k}.$$

Using (2), we have

$$\cos\theta = \frac{6 + 8 - 24}{(9 + 4 + 36)^{1/2}(4 + 16 + 16)^{1/2}} = \frac{-10}{42} = -0.2381.$$

The negative sign indicates that $\theta$ is between 90° and 180°, with $\cos(180° - \theta) = 0.2381$. This leads to $\theta = 180° - 76°14' = 103°46'$, approximately.

*The Scalar Product*

The expression in the numerator in (2) is given a special name. It is called the *scalar product* of $\mathbf{A}_1$ and $\mathbf{A}_2$, and denoted by $\mathbf{A}_1 \cdot \mathbf{A}_2$:

$$\mathbf{A}_1 \cdot \mathbf{A}_2 = a_1a_2 + b_1b_2 + c_1c_2. \tag{3}$$

From (2) we see that

$$\mathbf{A}_1 \cdot \mathbf{A}_2 = |\mathbf{A}_1||\mathbf{A}_2| \cos\theta. \tag{4}$$

The adjective "scalar" is used because the product $\mathbf{A}_1 \cdot \mathbf{A}_2$ is not a *vector*, but a *number*.

The scalar product of $\mathbf{O}$ and any vector is 0, because $|\mathbf{O}| = 0$. But if neither vector is $\mathbf{O}$, then $\mathbf{A}_1 \cdot \mathbf{A}_2 = 0$ if and only if $\cos\theta = 0$, i.e., if and only if the vectors are perpendicular.

We note that, in the case of the scalar product of a vector with itself,

$$\mathbf{A} \cdot \mathbf{A} = |\mathbf{A}|^2. \tag{5}$$

The following rules concerning scalar products are easily verifiable from (3):

$$\mathbf{A}_1 \cdot \mathbf{A}_2 = \mathbf{A}_2 \cdot \mathbf{A}_1, \tag{6}$$

$$(c\mathbf{A}) \cdot \mathbf{B} = c(\mathbf{A} \cdot \mathbf{B}), \tag{7}$$

$$(\mathbf{A}_1 + \mathbf{A}_2) \cdot \mathbf{B} = \mathbf{A}_1 \cdot \mathbf{B} + \mathbf{A}_2 \cdot \mathbf{B}. \tag{8}$$

The scalar product is useful in dealing with various geometric problems, as we shall see later on in this chapter. Here is one example.

**Example 2:** Let $\mathbf{A}$ be a nonzero vector, and let $\mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ be an arbitrarily selected vector. We require the expression of $\mathbf{R}$ as the sum of two vectors, one of them a multiple of $\mathbf{A}$ and the other either $\mathbf{O}$ or at right angles to $\mathbf{A}$ (see Fig. 18-5).

To solve this problem we write $\mathbf{R} = \mathbf{R}_1 + \mathbf{R}_2$ (notation as in Fig. 18-5) and concentrate on finding $\mathbf{R}_1$. Once $\mathbf{R}_1$ is found, we have $\mathbf{R}_2 = \mathbf{R} - \mathbf{R}_1$. Now, $\mathbf{R}_1$ is some multiple of $\mathbf{A}$, say $\mathbf{R}_1 = k\mathbf{A}$, and so $\mathbf{R} = k\mathbf{A} + \mathbf{R}_2$. Therefore, using (7) and (8), we have

$$\mathbf{R}\cdot\mathbf{A} = k\mathbf{A}\cdot\mathbf{A} + \mathbf{R}_2\cdot\mathbf{A}.$$

But $\mathbf{R}_2\cdot\mathbf{A} = 0$, as a result of the statement of the problem. Therefore $\mathbf{R}\cdot\mathbf{A} = k\mathbf{A}\cdot\mathbf{A} = k|\mathbf{A}|^2$, and hence

$$k = \frac{\mathbf{R}\cdot\mathbf{A}}{|\mathbf{A}|^2}, \qquad \mathbf{R}_1 = \frac{\mathbf{R}\cdot\mathbf{A}}{|\mathbf{A}|^2}\mathbf{A}. \tag{9}$$

The problem is thus solved.

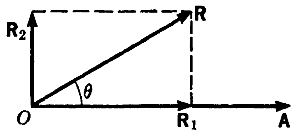



Fig. 18-5

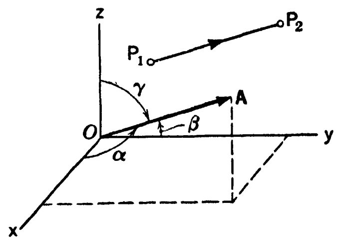



Fig. 18-6

As a numerical example, suppose

$$\mathbf{A} = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k}, \qquad \mathbf{R} = -\mathbf{i} + 5\mathbf{j} + 4\mathbf{k}.$$

Then $\qquad |\mathbf{A}|^2 = 14, \qquad \mathbf{R}\cdot\mathbf{A} = -2 + 15 - 4 = 9,$

and $\qquad \mathbf{R}_1 = \frac{9}{14}\mathbf{A}.$

### *Direction Cosines*

Assume a line determined by two distinct points $P_1$, $P_2$, and let the positive sense along the line be that from $P_1$ to $P_2$. This gives us what is called a *directed line.* Now consider a vector **A** issuing from $O$, parallel to and in the same sense as the directed segment $P_1P_2$ (see Fig. 18-6). Let $\alpha$, $\beta$, $\gamma$ be the angles which this vector makes with the positive axes of $x$, $y$, $z$, respectively. These angles are called the *direction angles* of the given directed line, and $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are called the *direction cosines* of the line.

If **A** is taken to be a unit vector (i.e., a vector of unit length), the direction cosines are the components of **A**, that is,

$$\mathbf{A} = \cos\alpha\mathbf{i} + \cos\beta\mathbf{j} + \cos\gamma\mathbf{k}. \tag{10}$$

This is merely a special case of the general fact that if $\mathbf{A} = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}$, then

$$a_1 = \mathbf{A}\cdot\mathbf{i}, \qquad b_1 = \mathbf{A}\cdot\mathbf{j}, \qquad c_1 = \mathbf{A}\cdot\mathbf{k}.$$

If $P_1$ is the point $(x_1, y_1, z_1)$ and $P_2$ is $(x_2, y_2, z_2)$, and if the distance $P_1P_2$ is $d$, then

$$\cos\alpha = \frac{x_2 - x_1}{d}, \qquad \cos\beta = \frac{y_2 - y_1}{d}, \qquad \cos\gamma = \frac{z_2 - z_1}{d}. \tag{11}$$

The vector

$$(x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k} \tag{12}$$

is parallel to and of the same length and sense as the directed line $P_1P_2$. In fact, this vector is $d$ times the unit vector $\mathbf{A}$ in Fig. 18-6.

**Example 3:** Find the direction cosines of the directed line from $(0, 1, \sqrt{2})$ to $(1, 0, 0)$.

In this case the vector (12) is $\mathbf{i} - \mathbf{j} - \sqrt{2}\,\mathbf{k}$, and the direction cosines are

$$\cos\alpha = \frac{1}{2}, \qquad \cos\beta = -\frac{1}{2}, \qquad \cos\gamma = -\frac{\sqrt{2}}{2}.$$

This indicates that $\alpha = 60°$, $\beta = 120°$, $\gamma = 135°$.

*Direction Components*

In many situations it is not necessary to deal with *directed* lines—that is, there is no need to assign a positive sense along the line. In such cases we usually specify the direction of the line merely by saying that it is parallel to some definite nonzero vector. The components of this vector are then called *direction components of the line*. Since the line is also parallel to every nonzero multiple of the vector, it follows that once we have a set of three direction components of the line, any proportional set also determines the direction of the line. If $l$, $m$, $n$ are direction components, we refer to the direction itself as "the direction $l{:}m{:}n$."

**Example 4:** The line through $(4, -6, 5)$ and $(-2, 6, -7)$ has the direction $6{:}-12{:}12$, which is the same as the direction $1{:}-2{:}2$, and also the same as $-1{:}2{:}-2$.

Corresponding to a given direction, there are two possible sets of direction cosines, one set being the negative of the other. To get direction cosines from $l{:}m{:}n$, we want a unit vector parallel to $l\mathbf{i} + m\mathbf{j} + n\mathbf{k}$. Hence we divide this vector by its own length, which is

$$(l^2 + m^2 + n^2)^{1/2}.$$

Thus, one set of direction cosines is

$$\frac{l}{(l^2 + m^2 + n^2)^{1/2}}, \qquad \frac{m}{(l^2 + m^2 + n^2)^{1/2}}, \qquad \frac{n}{(l^2 + m^2 + n^2)^{1/2}}.$$

The negatives of these form another set.

If two lines have directions $l_1{:}m_1{:}n_1$ and $l_2{:}m_2{:}n_2$, respectively, it is clear that the lines are parallel if and only if there is some nonzero constant $k$

such that $l_2 = kl_1$, $m_2 = km_1$, $n_2 = kn_1$. The lines are perpendicular if and only if

$$l_1l_2 + m_1m_2 + n_1n_2 = 0. \tag{13}$$

The student should make sure that he understands why this is so.

## EXERCISES

**1.** Find the cosine of the angle at the first mentioned vertex in each of the following triangles.
(a) (2, 4, 2), (4, 5, 4), (4, 6, 1).
(b) (2, 3, −1), (6, 4, 1), (5, 6, 4).
(c) (5, 4, 0), (3, 4, 1), (4, 6, −7).
(d) (4, 5, −3), (6, 9, 5), (8, 3, 5).

**2.** If the three direction angles $\alpha$, $\beta$, $\gamma$ for a line are acute and equal, what is the common value of the angles?

**3.** Using directions, determine which of the following sets of three points are collinear.
(a) (2, −1, 3), (5, 1, 2), (−1, −3, 4).
(b) (8, 3, 1), (−4, −5, 5), (2, −1, 2).
(c) (0, 2, −6), (3, 5, 0), (9, 11, 14).
(d) (1, −2, 4), (6, 1, 2), (−4, −5, 6).

**4.** Show by directions that certain of the following triangles are right triangles, and find the right angle. The points listed are vertices.
(a) (7, 3, 4), (1, 0, 6), (4, 5, −2).
(b) (4, 5, −6), (3, 6, −2), (2, 4, −4).
(c) (5, 6, 5), (−1, 3, 7), (2, 8, 0).
(d) (2, 4, 3), (4, 1, 9), (10, −1, 6).

**5.** Show that the four points (3, 4, 2), (5, 6, 1), (4, 8, 3), and (2, 6, 4) are the vertices of a square.

**6.** A directed line segment $P_1P_2$ has length 6 and it has the same direction as the vector $-2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$. If $P_1$ is $(-3, 2, 5)$, find $P_2$.

**7.** Two lines $L_1$, $L_2$ have directions 1:1:0 and 0:1:−1, respectively. Find a unit vector which is perpendicular to $L_1$ and makes an angle of 30° with $L_2$.

**8.** If $A$ is the point (4, 3, 6), find the point $P$ on the line through $O$ and (6, 2, 1) such that $PA$ and $OP$ are perpendicular.

**9.** Express $2\mathbf{i} + 5\mathbf{j} - 4\mathbf{k}$ as a multiple of $\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$ plus a vector perpendicular to the latter vector.

**10.** Suppose $A$ is (3, 0, 0), $B$ is (0, 4, 0), and $C$ is $(0, 0, c)$, where $c > 0$. What is $c$ if angle $ACB$ is 60°?

**11.** A room is 24 feet long, 16 feet wide, and 8 feet high. A line is drawn from each corner of the ceiling at one end of the room to the diagonally opposite corner of the floor at the other end of the room. Find the acute angle of intersection of these diagonals.

**12.** (a) Let $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$ be nonzero vectors, no one of which is in the plane of the other two. Let

$$\mathbf{B}_1 = \mathbf{A}_1, \qquad \mathbf{C}_1 = \frac{\mathbf{B}_1}{|\mathbf{B}_1|},$$

$$\mathbf{B}_2 = \mathbf{A}_2 - (\mathbf{A}_2\cdot\mathbf{C}_1)\mathbf{C}_1, \qquad \mathbf{C}_2 = \frac{\mathbf{B}_2}{|\mathbf{B}_2|},$$

$$\mathbf{B}_3 = \mathbf{A}_3 - (\mathbf{A}_3\cdot\mathbf{C}_1)\mathbf{C}_1 - (\mathbf{A}_3\cdot\mathbf{C}_2)\mathbf{C}_2, \qquad \mathbf{C}_3 = \frac{\mathbf{B}_3}{|\mathbf{B}_3|}.$$

Show that $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ are mutually perpendicular unit vectors.
(b) Calculate the $\mathbf{C}$ vectors if $\mathbf{A}_1 = 2\mathbf{i}$, $\mathbf{A}_2 = 3\mathbf{i} + 4\mathbf{j}$, $\mathbf{A}_3 = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.

## 18-3 Planes and Linear Equations

Our basic way of thinking about a plane is the following. A plane $M$ is uniquely determined if we know a point $P_0$ on $M$ and the direction of a line $L$ through $P_0$ perpendicular to $M$. There is a unique plane which goes through $P_0$ and is perpendicular to $L$. The condition that a point $P$ other than $P_0$ shall be in this plane is that the line through $P_0$ and $P$ shall be perpendicular to $L$ (see Fig. 18-7). There are also other geometrical conditions for determining a plane, but we consider this as our starting point in the discussion of planes.

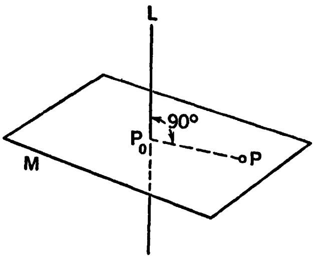


**Fig. 18-7**

If two lines are perpendicular to the same plane, they are parallel, and have the same direction. This direction is called *the direction normal to the plane.*

*The Point-Direction Equation of a Plane*

Consider the plane $M$ through $P_0(x_0, y_0, z_0)$, with $a{:}b{:}c$ as the direction of its normal. If $P(x, y, z)$ is any other point of $M$, the direction of $P_0P$ is $(x - x_0){:}(y - y_0){:}(z - z_0)$. The condition that $P_0P$ be perpendicular to the line $L$ normal to $M$ at $P_0$ is that

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \tag{1}$$

Hence this equation expresses the condition that $P$ be on $M$. We call (1) *the point-direction equation* of the plane.

**Example 1:** Find the equation of the plane which is the perpendicular bisector of the line segment joining $(-3, 4, 0)$ and $(5, -4, 6)$.

The mid-point of the segment has coordinates

$$x = \frac{-3 + 5}{2} = 1, \qquad y = \frac{4 - 4}{2} = 0, \qquad z = \frac{0 + 6}{2} = 3.$$

The direction of the segment is

$$5 + 3 : -4 - 4 : 6 - 0, \quad \text{or} \quad 4 : -4 : 3.$$

Hence the equation of the required plane is

$$4(x - 1) - 4(y - 0) + 3(z - 3) = 0,$$

or

$$4x - 4y + 3z = 13.$$

The following theorem expresses a very important fact about planes.

THEOREM 18-A. *Every plane is characterized by a linear equation in $x$, $y$, $z$. Conversely, an equation $Ax + By + Cz + D = 0$, in which $A$, $B$, $C$ are not all zero, has a plane as its locus. The direction normal to the plane is $A:B:C$.*

*Proof.* We have seen that every plane can be characterized by a point-direction equation of the form (1). This equation is linear in $x$, $y$, $z$. Now consider the locus of the linear equation

$$Ax + By + Cz + D = 0. \tag{2}$$

We can always find at least one point on this locus by setting some two of the coordinates equal to zero and solving for the third. For example, if $C \neq 0$, we can set $x_0 = y_0 = 0$ and obtain $z_0 = -D/C$, yielding $(0, 0, -D/C)$ on the locus. Now, if $(x_0, y_0, z_0)$ is some definite point on the locus, we have $Ax_0 + By_0 + Cz_0 + D = 0$. In view of this, (2) is equivalent to

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

This, however, is the equation of the plane through $(x_0, y_0, z_0)$ with $A:B:C$ as its normal direction. Hence this plane is the locus of (2).

*The Distance from a Point to a Plane*

The distance $d$ from a point $P_0(x_0, y_0, z_0)$ to the plane $Ax + By + Cz + D = 0$ is

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{(A^2 + B^2 + C^2)^{1/2}}. \tag{3}$$

This formula can be worked out by a method exactly like the one used in deducing the formula for the distance from a point to a line in plane analytic geometry (see § 7-2) and Exercise 1 in § 7-1).

For an alternative method of getting a result equivalent to (3), using vectors, see Exercise 4.

**Example 2:** Find the distance (a) of the origin and (b) of the point (2, 1, 3) from the plane $2x - y + 2z - 6 = 0$. Are the points on the same or opposite sides of the plane?

Using (3) in case (a) we have

$$d = \frac{|6|}{\sqrt{9}} = 2,$$

while in case (b) we have

$$d = \frac{|4 - 1 + 6 - 6|}{\sqrt{9}} = 1.$$

The two points are on *opposite* sides of the origin. The reason for this is the following: The expression $2x - y + 2z - 6$ is equal to 0 when $(x, y, z)$ is *on* the plane. For points not on the plane, the expression is either positive or negative; those for which it is positive are all on one side of the plane, and the points for which the expression is negative are on the other side of the plane. In this instance the expression is negative at (0, 0, 0) and positive at (2, 1, 3).

*The Plane Through Three Given Points*

If three points are given, not all on one straight line, there is a unique plane that contains all three points. To find an equation of this plane, it suffices to determine coefficients $A$, $B$, $C$, $D$, not all zero, in such a way that, if $P_1$, $P_2$, $P_3$ are the three points, $P_k$ having coordinates $(x_k, y_k, z_k)$, then

$$\left.\begin{aligned} Ax_1 + By_1 + Cz_1 + D &= 0,\\ Ax_2 + By_2 + Cz_2 + D &= 0,\\ Ax_3 + By_3 + Cz_3 + D &= 0. \end{aligned}\right\} \quad (4)$$

The equation of the plane will then be

$$Ax + By + Cz + D = 0. \quad (5)$$

**Example 3:** Find the plane through (1, 2, 5), (−1, 2, 9), and (4, −4, −10). In this case the system (4) becomes

$$\begin{aligned} A + 2B + 5C + D &= 0,\\ -A + 2B + 9C + D &= 0,\\ 4A - 4B - 10C + D &= 0. \end{aligned}$$

We solve by elimination, seeking to express three of the coefficients in terms of the fourth, which is then assigned some arbitrary nonzero value. From the first and second equations we obtain

$$2A - 4C = 0, \quad \text{or} \quad A = 2C.$$

From the first and third, and the second and third, respectively, we obtain

$$\begin{aligned} -3A + 6B + 15C &= 0,\\ -5A + 6B + 19C &= 0. \end{aligned}$$

Elimination of $A$ now gives

$$12B + 18C = 0, \quad \text{or} \quad B = -\frac{3}{2}C.$$

Finally, $\quad D = -A - 2B - 5C = -2C + 3C - 5C = -4C.$

The final result, on setting $C = 2$, is the equation

$$4x - 3y + 2z - 8 = 0$$

for the required plane.

In a problem of this kind it may turn out that a certain one of the four coefficients is 0. The other coefficients cannot then be expressed in terms of that particular one.

An alternative solution of the general problem of a plane through three points can be made as follows: If $P_1$, $P_2$, $P_3$, and $P(x, y, z)$ are all on the plane, then equations (4) and (5) all hold simultaneously. This means that the four "unknowns" $A$, $B$, $C$, $D$ are not all zero and satisfy the four homogeneous linear equations whose coefficients form the fourth-order determinant

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix}. \tag{6}$$

By the extension of the results of § 17-4 concerning homogeneous linear equations, it must then be true that the determinant (6) is equal to 0. For, if it were not, Cramer's rule as applied to the system of four equations in the unknowns $A$, $B$, $C$, $D$, would yield the unique solution $A = B = C = D = 0$. Hence, setting the determinant in (6) equal to 0 gives a linear equation in $x$, $y$, $z$ which is the equation of the plane.

**Example 4:** Solving Example 3 by this method, we have

$$\begin{vmatrix} x & y & z & 1 \\ 1 & 2 & 5 & 1 \\ -1 & 2 & 9 & 1 \\ 4 & -4 & -10 & 1 \end{vmatrix} = 0.$$

Expanding by minors of the first row, we have

$$\begin{vmatrix} 2 & 5 & 1 \\ 2 & 9 & 1 \\ -4 & -10 & 1 \end{vmatrix} x - \begin{vmatrix} 1 & 5 & 1 \\ -1 & 9 & 1 \\ 4 & -10 & 1 \end{vmatrix} y$$
$$+ \begin{vmatrix} 1 & 2 & 1 \\ -1 & 2 & 1 \\ 4 & -4 & 1 \end{vmatrix} z - \begin{vmatrix} 1 & 2 & 5 \\ -1 & 2 & 9 \\ 4 & -4 & -10 \end{vmatrix} = 0.$$

Upon calculation of the determinants of third order, we obtain

$$24x - 18y + 12z - 48 = 0,$$

or

$$4x - 3y + 2z - 8 = 0.$$

This method is not presented as a recommendation for the use of determinants, but merely for illustration.

Several other types of problems involve the determination of a plane through the solution of three homogeneous equations in the coefficients $A$, $B$, $C$, $D$. Some such problems will be found in the exercises. Here is one type of condition leading to a homogeneous linear equation: If the plane being sought is perpendicular to a given plane whose normal direction is $l{:}m{:}n$, then $A{:}B{:}C$ and $l{:}m{:}n$ are perpendicular directions, and so

$$lA + mB + nC = 0.$$

## EXERCISES

**1.** Find the equation of the plane:
(a) Through the point $(-3, 2, 0)$ and perpendicular to the line through the points $(0, 2, 3)$, $(2, 1, 4)$;
(b) Through the point $(4, -5, 1)$ and parallel to the plane $2x + y - 5z = 0$;
(c) Parallel to the plane $3x - 12y + 4z = 24$, 2 units from $(0, 0, 0)$ and 3 units from $(3, 0, 1)$;
(d) Through the points $(1, 2, -3)$, $(-2, 3, 0)$ and perpendicular to the plane $x - y - 2z = 3$.

**2.** Find the equation of the plane:
(a) Tangent to the sphere $x^2 + y^2 + z^2 - 6x + 4y = 156$ at the point $(7, 1, 12)$;
(b) Which is the perpendicular bisector of the line segment from $(2, -4, 3)$ to $(-1, 2, 7)$;
(c) Perpendicular to the line through $(3, -2, -3)$ and $(1, 2, 1)$, 2 units from the origin, and 4 units from $(-2, 1, 1)$;
(d) Through the point $(-2, 1, 1)$ and perpendicular to each of the planes $x + 3y + z = 2$, $2x - y - z = 4$.

**3.** If $\mathbf{R}$ is the vector from $O$ to $P$, if $\mathbf{R}_0$ is the vector from $O$ to $P_0$, and if $\mathbf{N}$ is a nonzero vector, show that $P$ is on the plane through $P_0$ with normal direction parallel to $\mathbf{N}$ if and only if $\mathbf{N}\cdot(\mathbf{R} - \mathbf{R}_0) = 0$.

**4.** Suppose the plane $M$ does not go through $O$. Let $\mathbf{n}$ be the unit vector from $O$ perpendicularly toward $M$, and let $p$ be the distance from $O$ to $M$ (schematic diagram in Fig. 18-8). Show that, if $\mathbf{R}$ is the vector from $O$ to $P$, the condition that $P$ be on $M$ is $\mathbf{R}\cdot\mathbf{n} = p$. Hence, if $\mathbf{R}_0$ is the vector from $O$ to $P_0$ and $d$ is the distance from $P_0$ to $M$, show that $d = |\mathbf{R}_0\cdot\mathbf{n} - p|$. Compare this result with (3) and describe how to find $\mathbf{n}$ and $p$ in terms of the coefficients $A$, $B$, $C$, $D$. What changes must be made in all of this if $M$ *does* go through $O$?

**5.** Show that the planes $2x - y + 2z = 12$, $6x - 3y + 6z + 45 = 0$ are parallel, and find the perpendicular distance between them.

**6.** Why cannot $Ax + By + Cz + D$ be negative at some and positive at other points on the same side of the plane $Ax + By + Cz + D = 0$? *Suggestion:* The reason is related to Theorem 6-A.

**7.** Find the equations of the faces of the tetrahedron whose vertices are the points (0, 0, 0), (0, 0, 3), (2, 0, 1), (1, 2, 1).

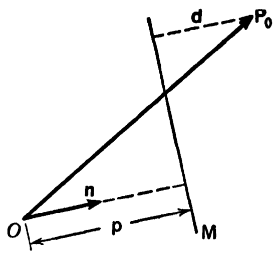



**Fig. 18-8**

**8.** Find the equation of the plane through the points (4, 1, 2), (5, 2, 3), (−3, 3, 1).

**9.** What is the acute angle between the planes $x + y + 4 = 0$, $2x + y - 2z = 3$?

**10.** Find an equation of the plane $M$ if the $y$-axis lies in $M$ and $M$ contains the point (2, 4, −3).

**11.** Find the shortest distance from the plane $12x + 4y + 3z = 327$ to the sphere $x^2 + y^2 + z^2 + 4x - 2y - 6z = 155$.

**12.** Find the equations of the planes which bisect the angles between the planes $12x - 5y = 39$, $3x + 4z = 24$. Verify that the planes thus found are perpendicular.

## 18-4 Planes and Straight Lines

In § 18-1 we have seen how to represent a straight line parametrically. If $(x_0, y_0, z_0)$ is on the line and if the line has direction $l{:}m{:}n$, parametric equations for the line are

$$x = x_0 + lt, \qquad y = y_0 + mt, \qquad z = z_0 + nt. \tag{1}$$

There are other ways of representing a line analytically. For example, if two planes intersect, their intersection is a straight line, and the equations of the planes, taken simultaneously, form a pair of equations describing the line. The direction of the line may be found by making use of the fact that the line, being in each plane, is perpendicular to the normal to each plane. When these facts are expressed algebraically, we have two homogeneous linear equations from which to find the direction of the line. The method of solution is explained in § 17-1 (see Theorem 17-A).

**Example 1:** Find the direction $l{:}m{:}n$ of the line described by the two equations

$$2x + 3y + 5z = 0,$$
$$x - y + 4z = 4. \tag{2}$$

The two normal directions here are 2:3:5 and 1:−1:4. The perpendicularity conditions are

$$2l + 3m + 5n = 0,$$
$$l - m + 4n = 0.$$

Solving as in Theorem 17-A, we find

$$l{:}m{:}n = \begin{vmatrix} 3 & 5 \\ -1 & 4 \end{vmatrix} : \begin{vmatrix} 5 & 2 \\ 4 & 1 \end{vmatrix} : \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = 17{:}-3{:}-5.$$

If a line is presented as the intersection of two planes, parametric equations for the line may be obtained, using one of the coordinates as a parameter. The coordinate selected must be such that the line is not perpendicular to the corresponding axis.

**Example 2:** Obtain parametric equations of the line in Example 1, taking $z$ as the parameter.

We solve equations (2) for $x$ and $y$ in terms of $z$, by elimination. The results are

$$x = -\frac{17}{5}z + \frac{12}{5}, \qquad y = \frac{3}{5}z - \frac{8}{5}. \tag{3}$$

These two equations, together with $z = z$, form a set of parametric equations of the line. Each equation in (3) is linear, and therefore represents a plane. The first equation does not contain $y$. It represents a plane parallel to the $y$-axis. The second equation represents a plane parallel to the $x$-axis.

*Symmetric Equations of a Line*

If the parametric equations (1) are written in the form

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}, \tag{4}$$

these equations are called *symmetric equations* of the line.

Equations (3) can easily be rewritten in the symmetric form (4):

$$\frac{x - \frac{12}{5}}{-17} = \frac{y + \frac{8}{5}}{3} = \frac{z}{5}.$$

The symmetric form is not unique, because any point $(x_0, y_0, z_0)$ on the line can be used, and $l{:}m{:}n$ can be replaced by any proportional set.

*Simultaneous Linear Equations*

The algebraic problem of solving simultaneous linear equations in $x$, $y$, $z$ is clarified and made readily understandable if it is viewed from the geometric point of view.

A linear equation represents a plane. A system of two linear equations (in $x$, $y$, $z$) represents two planes. To solve the system means to find all points $(x, y, z)$ which lie on both planes. There are three cases: (1) the planes intersect in a line; (2) the planes are identical (i.e., one equation is a multiple of the other); (3) the planes are distinct and parallel. In cases 1 and 2 the algebraic problem has a solution, but it is not unique. In case 3 there is no solution, because the planes have no point in common.

A system of three linear equations represents three planes. Here there are many possibilities. We shall discuss the situation in relation to the considerations in § 17-4. Now we use $x$, $y$, $z$ in place of $x_1$, $x_2$, $x_3$. Suppose the system is

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1, \\ a_{21}x + a_{22}y + a_{23}z &= b_2, \\ a_{31}x + a_{32}y + a_{33}z &= b_3. \end{aligned} \tag{5}$$

Let $D$ be the determinant of the system.

*Case 1:* $D \neq 0$. In this case the three planes are all distinct and they have a single common point. The algebraic system has a unique solution.

*Case 2:* $D = 0$, but some element has a minor $\neq 0$. In this case two of the planes are distinct and intersect in a line. The third plane either goes through this same line, in which case the system has a solution, but not a unique one, or else the third plane is parallel to the line of intersection of the first two planes, in which case there is no common point, and the algebraic system is inconsistent.

*Case 3:* $D = 0$ and the minor of each element is 0. In this case all three planes have the same normal direction. They may be all distinct and parallel, or there may be just two distinct parallel planes, or all three equations may define the same plane. In this last situation the system has a solution (not unique). In the other situations there is no common point and the algebraic system is inconsistent.

If the system (5) is homogeneous, this means that $b_1 = b_2 = b_3 = 0$. In this case all three planes pass through the origin. If $b_1 \neq 0$, the equation with $b_1$ replaced by 0 represents a plane parallel to the original one, but through the origin.

Referring back now to Examples 2 and 3 in § 17-4, we can interpret the results of those examples in geometric terms. Example 2 represents three distinct planes which intersect along the line through (0, 0, 0) and (9, −4, −6). Example 3 represents one single plane; the three equations are identical except for different scalar multiples.

## EXERCISES

**1.** A straight line through the point $(2, -3, 1)$ makes angles $60°$, $45°$, $60°$ with the $x$-, $y$-, and $z$-axes, respectively. (a) Write symmetric equations for the line. (b) What is the equation of the plane through the line and parallel to the $y$-axis?

**2.** (a) Write a set of symmetric equations of the line determined by the equations $4x - 3y - z = 1$, $2x + 4y + z = 5$. (b) What is the direction of the line? (c) Where does the line pierce the plane $x + y + z = 32$?

**3.** Find planes through the line $3x + 2y - z = 4$, $x + 3y + z = 5$ and (a) parallel to the $x$-axis; (b) parallel to the $y$-axis; (c) parallel to the $z$-axis.

**4.** Find the line (a) through $(1, 2, 0)$ and perpendicular to the plane $3x - y - 2z = 2$; (b) through $(4, -1, 3)$ and parallel to the line $5x - 3y = 41$, $7y + 5z + 14 = 0$.

**5.** Find the line (a) through $(3, -1, 2)$ and perpendicular to the plane $2x - 3y + 4z = 4$; (b) through $(0, 2, -3)$ and parallel to the line $y = 2x - 7$, $z = 3x + 4$.

**6.** Find the plane through $(1, 3, 5)$ and parallel to the plane of the two vectors $3\mathbf{i} - \mathbf{j} + 2\mathbf{k}$, $4\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$.

**7.** Find the plane through $(2, 4, 3)$ and parallel to the lines whose equations are $2x - 5y + 11z = 4$, $x - 6y + z = -3$, and

$$\frac{x-4}{7} = \frac{y-3}{-2} = \frac{z+1}{5}.$$

**8.** (a) If the equations of two distinct but intersecting planes are denoted by $f(x, y, z) = 0$, $g(x, y, z) = 0$, explain why each but one of the planes through the line of intersection of these planes has an equation of the form $f(x, y, z) + kg(x, y, z) = 0$, where $k$ is some constant. Which plane is not so represented? (b) Use the idea suggested in (a) to find the plane through the line $3x - 2y - z = 1$, $4x + 3y - z = 4$ and the point $(1, 1, 1)$.

**9.** Find the plane through $(2, 3, 1)$ and the intersection of the planes $x + 4y - z = 3$, $2x + 3y - z = 4$. See Exercise 8(a).

**10.** Find the plane through the line $4x - y - 3z = 4$, $5x - 2y + z = 7$ and perpendicular to the $xz$-plane. See Exercise 8(a).

**11.** Find the plane through the line $4x + z = 1$, $4y + z = 5$ and perpendicular to the plane $x + 2y - z = 3$. See Exercise 8(a).

**12.** (a) Find the point of intersection of the line $x - 2y + 3 = 0$, $2x - 2y - z + 3 = 0$ and the line through the points $(-1, 16, -4)$, $(3, -12, 6)$. (b) At what angles do the lines intersect? (c) Find the equation of the plane determined by the two lines.

**13.** Find a point on the line joining $(2, 1, -2)$ and $(1, -3, 2)$ such that it is equidistant from $(0, 1, 1)$ and $(1, 2, 3)$.

14. Find the length of the projection of the straight line joining the points (2, 1, 3), (3, 3, 5) upon the straight line determined by the points (1, 2, 1), (4, −1, 3).

15. Find the center and radius of the circle cut from the sphere $x^2 + y^2 + z^2 + 2x - 4y - 4z = 16$ by the plane $3x - 4y - 12z = 17$.

## 18-5 The Cross Product of Two Vectors

If **A** and **B** are vectors, we define what is called the *cross product* $\mathbf{A} \times \mathbf{B}$ as follows:

$$\mathbf{O} \times \mathbf{B} = \mathbf{A} \times \mathbf{O} = \mathbf{O}, \quad \text{by definition;} \tag{1}$$

if neither **A** nor **B** is **O**, and if $\theta$ is the angle between **A** and **B**, then

$$\mathbf{A} \times \mathbf{B} = \mathbf{C},$$

where **C** is the vector perpendicular to the plane of **A** and **B**, of magnitude

$$|\mathbf{C}| = |\mathbf{A}|\,|\mathbf{B}| \sin\theta, \tag{2}$$

the sense of **C** being such that **A**, **B**, **C** in that order form a right-handed system (see Fig. 18-9). This means that as the plane of **A** and **B** is viewed from the tip of **C**, the angle $\theta$ from **A** to **B** is generated by a counterclockwise rotation.

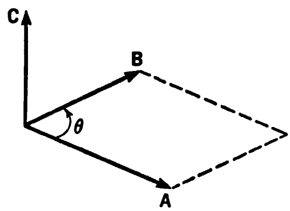


**Fig. 18-9**

Our main interest in cross products just now arises from the fact that if **A** and **B** have components $a_1$, $a_2$, $a_3$ and $b_1$, $b_2$, $b_3$, respectively, then $\mathbf{A} \times \mathbf{B}$ has the components

$$\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \quad \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}. \tag{3}$$

In view of earlier work (see Theorem 17-A in § 17-1 and Example 1 in § 18-4), we recognize that the three second-order determinants in (3) are components of a vector perpendicular to both **A** and **B**. But now we are also describing how the length and sense of this vector depend upon **A** and **B**. The actual proof that $\mathbf{A} \times \mathbf{B}$ has the components (3) requires more space than is available for the brief treatment in this book. The method consists in expanding the product

$$(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})$$

with the aid of the rules

$$(c\mathbf{A}) \times \mathbf{B} = c(\mathbf{A} \times \mathbf{B}), \tag{4}$$

$$\mathbf{A}_1 \times (\mathbf{A}_2 + \mathbf{A}_3) = \mathbf{A}_1 \times \mathbf{A}_2 + \mathbf{A}_1 \times \mathbf{A}_3, \tag{5}$$

and then working out the particular products $\mathbf{i} \times \mathbf{i} = \mathbf{O}$, $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, and so on. We shall not insist on the details, which are usually considered fully in books on vector analysis.

Now suppose that **A**, **B**, **C** are any three vectors, with components $(a_1, a_2, a_3)$, and so on. According to the foregoing,

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} + a_2\begin{vmatrix} b_3 & b_1 \\ c_3 & c_1 \end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}. \tag{6}$$

But

$$\begin{vmatrix} b_3 & b_1 \\ c_3 & c_1 \end{vmatrix} = -\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix},$$

and so we see that

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \tag{7}$$

Formula (6) is obtained when the determinant in (7) is expanded by minors of the first row.

There is a very nice geometric interpretation of $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$. The absolute value of this expression is equal to the volume of the parallelepiped formed with edges along **A**, **B**, **C**, as in Fig. 18-10. This is clear as soon as we observe that $\mathbf{B}\times\mathbf{C}$ is a vector of magnitude equal to the area of the parallelogram formed on **B** and **C** as adjacent sides, and then recall the geometric meaning of the scalar product [see (4) in § 18-2]. Whether $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$ is equal to this volume or its negative depends on whether or not **A**, **B**, **C** form a right-handed system.

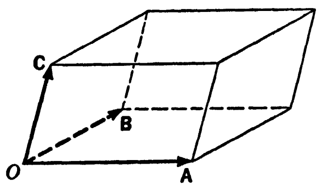


**Fig. 18-10**

If it happens that one of the vectors is **O**, or in the plane of the other two vectors, the volume reduces to 0. We have, then, a geometric interpretation of the condition for the determinant (7) to be 0. Its rows, regarded as vectors, must be linearly dependent; that is, some row must be a linear combination of the other two.

**Example:** Find the volume of the parallelepiped, three coterminous edges of which are from (1, 2, 5) to (2, −1, 7), (3, 3, 4), and (0, 5, 1), respectively.

By translation, the parallelepiped in question is congruent to one with coterminous edges from (0, 0, 0) to (1, −3, 2), (2, 1, −1), and (−1, 3, −4), respectively. Hence, the required volume is the absolute value of

$$\begin{vmatrix} 1 & -3 & 2 \\ 2 & 1 & -1 \\ -1 & 3 & -4 \end{vmatrix}.$$

This works out to be −14, so the required volume is 14 cubic units.

EXERCISES

1. Show that $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \mathbf{B}\cdot(\mathbf{C}\times\mathbf{A}) = \mathbf{C}\cdot(\mathbf{A}\times\mathbf{B})$.
2. If **A**, **B**, **C** are vectors from the origin to three noncollinear points $P$, $Q$, $R$, respectively, show that the vector $\mathbf{A}\times\mathbf{B}+\mathbf{B}\times\mathbf{C}+\mathbf{C}\times\mathbf{A}$ is either **O** or perpendicular to the plane of $PQR$.
3. Use scalar and cross products to express the condition that the plane through two vectors **A**, **B** be perpendicular to the plane through two vectors **C**, **D**.
4. Find the volume of the parallelepiped, three coterminous edges of which:
 (a) join the origin to $(2, -3, 5)$, $(-1, 4, 2)$, $(2, 3, 0)$;
 (b) join $(2, 1, 4)$ to $(4, 4, 5)$, $(3, 3, 9)$, $(0, 5, 7)$.
5. Consider the vectors $\mathbf{A} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$, $\mathbf{B} = 2\mathbf{i} + \mathbf{j} - \mathbf{k}$, $\mathbf{C} = \mathbf{i} - \mathbf{j} + 3\mathbf{k}$ with terminal points $P$, $Q$, $R$, respectively. Use a cross product to find a unit vector perpendicular to the plane of $PQR$, and then project **A** on it to find the distance from the origin to the plane.
6. Suppose $\mathbf{A} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{B} = 3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$, $\mathbf{C} = -\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}$, $\mathbf{D} = -\mathbf{i} - 2\mathbf{j} + \mathbf{k}$. Let $L_1$ be the line through the tips of **A** and **B**, $L_2$ the line through the tips of **C** and **D**. Use a cross product to find a unit vector **n** perpendicular to each of the lines $L_1$, $L_2$. If **V** is a vector having the length and direction of the line segment joining the tips of **A** and **C**, explain why $|\mathbf{V}\cdot\mathbf{n}|$ is the perpendicular distance between $L_1$ and $L_2$. Find this distance.

## 18-6 Surfaces in Space

To begin with, we consider surfaces of especially simple kinds, and examine the ways in which we can identify the surfaces or gain information about them by looking at their equations.

*Cylinders*

Consider a curve $C$ lying in a plane $M$. Select a straight line $L$ which cuts $M$ at a single point. Through each point of $C$ draw a straight line parallel to $L$. All these latter lines, taken together, form a configuration called a cylindrical surface. These lines are called *elements* of the cylinder. The curve $C$ is called a directrix of the cylinder. If $C$ is a circle and if $L$ is perpendicular to $M$, we call the cylinder a *right circular cylinder*. An oblique plane section of such a cylinder is an ellipse. There are, of course, parabolic cylinders, hyperbolic cylinders, and so on.

When the elements of a cylinder are parallel to a coordinate axis, the cylinder is described by an equation which does not involve the corresponding coordinate. (This has already been mentioned in § 6-6.) Thus, if $f(x, y) = 0$ is the equation of a certain curve $C$ in the $xy$-plane (when the point of view is that of *plane* geometry), then, from the point of view of three-dimensional geometry, $f(x, y) = 0$ is the equation of the cylinder with directrix $C$ and elements parallel to the $z$-axis.

**Example 1:** The equation $x^2 + z = 4$ defines a parabolic cylinder with elements parallel to the $y$-axis. The equation $y = x^2$ defines a parabolic cylinder with elements parallel to the $z$-axis. Figure 18-11 depicts parts of these surfaces in the first octant, and shows how they intersect.

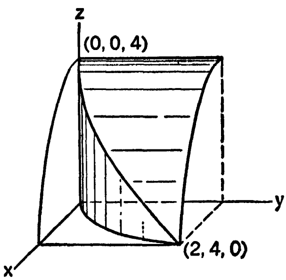



**Fig. 18-11**

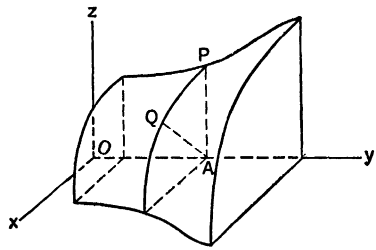



**Fig. 18-12**

*Surfaces of Revolution*

If a surface has an axis of symmetry and if all plane sections of the surface at right angles to this axis are circles, the surface is called a *surface of revolution.* Here is one way to generate a surface of revolution: Take a curve in the $xy$-plane defined by an equation $y = f(x)$, where $f$ is continuous when $a \leq x \leq b$. Revolve the plane about the $x$-axis. Then the curve will generate a surface of revolution with the $x$-axis as its axis of symmetry. In like fashion we can obtain surfaces with the $y$-axis or $z$-axis as axis of symmetry.

In any of these cases it is easy to pass from the equation of the plane curve to the equation of the surface of revolution which it generates. Suppose, for instance, that we are revolving a curve about the $y$-axis (see Fig. 18-12). The equation of the curve in the $yz$-plane expresses $AP$ as a function of $OA$, say $AP = f(y)$, where $y = OA$. Now, on the surface of revolution, $AQ = AP$. Hence, if $Q$ has coordinates $(x, y, z)$, we have $AQ = \sqrt{x^2 + z^2} = f(y)$. Thus the equation of the surface of revolution is $x^2 + z^2 = [f(y)]^2$, where the equation of the generating curve was $z = f(y)$.

**Example 2: If the line $\frac{y}{3} + \frac{z}{2} = 1$ in the $yz$-plane is revolved about the $y$-axis, the resulting surface of revolution is a right circular cone (of two nappes) with vertex at $y = 3$ on the $y$-axis. To get the equation of the cone we replace $z$ by $\sqrt{x^2 + z^2}$, solve for the radical, and then rationalize by squaring:**

$$\frac{y}{3} + \frac{\sqrt{x^2 + z^2}}{2} = 1, \qquad \sqrt{x^2 + z^2} = 2\left(1 - \frac{y}{3}\right),$$

$$x^2 + z^2 = \frac{4}{9}(y - 3)^2.$$

If we had not performed the squaring, we would have had the equation of just one nappe of the cone—the part on which $y \leq 3$.

*The Standard Quadric Surfaces*

A quadric surface is any surface whose equation in rectangular coordinates is of the second degree. The cylinders of Example 1 and the cone of Example 2 are quadric surfaces. We are now going to consider *ellipsoids*, two kinds of *paraboloids*, and two kinds of *hyperboloids*, so placed in relation to the coordinate axes that they have comparatively simple equations.

**Example 3:** The equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ defines what is called an ellipsoid. If $a = b = c$, we have a sphere as a special case. If $a = b \neq c$, we have an ellipsoid of revolution with circular sections in planes perpendicular to the $z$-axis. Figure 18-13 shows an ellipsoid, with the first octant portion cut away to give a better idea of the shape of the surface.

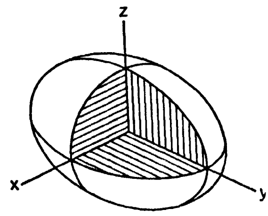



**Fig. 18-13**

**Example 4:** The equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = z$ defines what is called an *elliptic paraboloid*. If $a = b$ it is a paraboloid of revolution about the $z$-axis. Plane sections of the surface at right angles to the $z$-axis are ellipses. Plane sections parallel to the $z$-axis are parabolas. For a representation see Fig. 18-14.

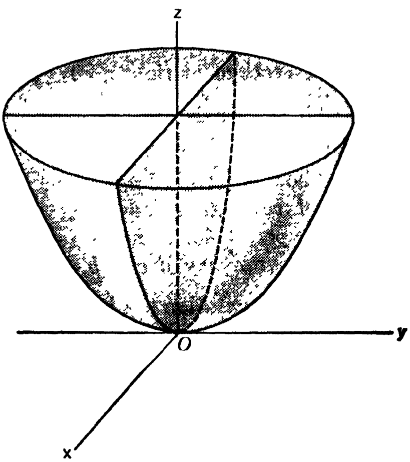



**Fig. 18-14**

**Example 5:** The equation $\frac{x^2}{a^2} - \frac{y^2}{b^2} = -z$ defines what is called a *hyperbolic paraboloid.*

This surface is saddle-shaped in the vicinity of the origin. A person sitting erect in the saddle would be astride the $y$-axis, with his body along the positive $z$-axis. Planes $z =$ constant cut the surface in hyperbolas. Planes parallel to the $z$-axis generally cut the surface in parabolas, but in the special cases of planes parallel to $x/a = y/b$ or $x/a = -y/b$, the intersections are straight lines on the surface. The surface can be built up entirely from these lines, and is on that account called a *ruled surface.* See Fig. 18-15. Interesting string models of hyperbolic paraboloids can be constructed.

**Example 6:** There are two kinds of hyperboloids. If the hyperbola $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$ in the $xz$-plane is revolved about the $z$-axis, it generates a surface

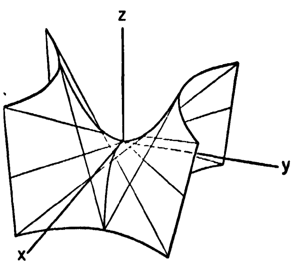


**Fig. 18-15**

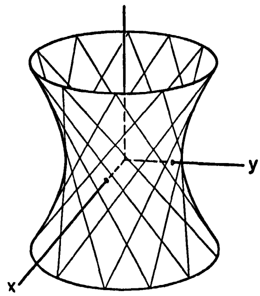


**Fig. 18-16**

called a *hyperboloid of one sheet,* whose equation is $\frac{x^2 + y^2}{a^2} - \frac{z^2}{c^2} = 1$. A more general type of hyperboloid of one sheet is represented by the equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$. All plane sections at right angles to the $z$-axis are ellipses, and plane sections through the $z$-axis are hyperbolas. See Fig. 18-16. This surface also can be built up out of straight lines.

**Example 7:** If a hyperbola is revolved about the axis through the foci, it generates a hyperboloid of revolution consisting of two separated parts. It is called a *two-sheeted* hyperboloid. A more general hyperboloid of this type, not having rotational symmetry if $a \neq b$, is defined by the equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$. For a representation see Fig. 18-17.

## EXERCISES

**1.** Describe and sketch the surface represented by each equation. If it is a cylinder, state the direction of its elements. If it is a surface of revolution, name the axis of revolution.

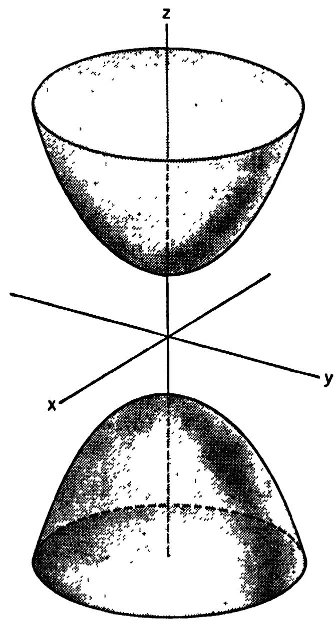



**Fig. 18-17**

(a) $9x^2 + 16y^2 = 144.$
(b) $x^2 + z^2 = 4y.$
(c) $x^2 + z^2 = 2z.$
(d) $x^2 + y^2 + 4z = 4.$
(e) $z^2 + 6 = 5z.$
(f) $4(x^2 + y^2) = (z - 5)^2.$
(g) $z^2 - 4z = 2y.$
(h) $y^2 + x = 4.$

**2.** Proceed as directed in Exercise 1.
(a) $z^2 - x^2 = 4.$
(b) $y^2 = x^2 + z^2.$
(c) $yz = 4.$
(d) $9(x^2 + z^2) + 16y^2 = 144.$
(e) $y^2 + 3x = 9.$
(f) $z^2 = 4(x^2 + y^2).$
(g) $16 - y^2 = 8z.$
(h) $4 - z = x^2 + y^2.$

**3.** Draw the solid lying in the first octant and bounded by the coordinate planes and the surfaces $3z = 9 - x^2$, $x + 2y = 6$.

**4.** Draw a figure showing the surfaces $16x^2 + 9y^2 = 144$, $25x^2 + 9z^2 = 225$ and their intersection in the first octant. Find the plane in which this curve of intersection lies.

**5.** Draw a figure showing the surfaces $z = 4 - x^2$, $4x^2 + (y - 4)^2 = 16$, and their intersection. What cylinder with elements parallel to the $x$-axis passes through this curve of intersection?

**6.** Identify the surfaces $x^2 + z^2 = 2y$, $y = z$, and sketch their intersection. Find a cylinder parallel to the $z$-axis through their intersections, and hence identify by name the type of the curve of intersection.

**7.** Find the equation of the surface of revolution generated by revolving the given plane curve about the axis mentioned. Name the surface if it is of a type heretofore discussed.

(a) $x^2 = 4y$, about the $y$-axis.
(b) $y^2 = 2x$, about the $y$-axis.
(c) $2y + 3z = 6$, about the $z$-axis.
(d) $9x^2 - 4z^2 = 36$, about the $z$-axis.
(e) $16x^2 - 9y^2 = 144$, about the $x$-axis.
(f) $4z^2 + (y - 4)^2 = 16$, about the $y$-axis.

**8.** Describe each surface, name it, and make a rough perspective sketch of the surface. Make a separate set of diagrams showing the way the surface intersects the coordinate planes, if it does intersect. (Also, plane sections at right angles to an axis may be useful as an aid in visualizing the surface.)

(a) $\frac{x^2}{16} + \frac{y^2}{9} + \frac{z^2}{25} = 1.$

(b) $-\frac{x^2}{16} + \frac{y^2}{9} + \frac{z^2}{25} = 1.$

(c) $\frac{x^2}{16} - \frac{y^2}{9} - \frac{z^2}{25} = 1.$

(d) $36z = 144 - 9x^2 - 4y^2.$

(e) $4x^2 - 9y^2 + 36z^2 = -36.$

(f) $4x^2 - 9y^2 + 36z^2 = 0.$

**9.** (a) What form does the equation $z = xy$ take if new axes $X$, $Y$, $Z$ are taken, with the $Z$-axis the same as the $z$-axis, but the $XY$-axes turned 45° relative to the $xy$-axes? (b) What is the name of the surface? (c) What are the plane sections of the surface by planes $z$ = constant?

## 18-7 Curves in Space

To describe a curve in space analytically, the natural general method is that of parametric representation; the coordinates $(x, y, z)$ of a point on the curve are expressed as functions of some auxiliary variable, called a parameter:

$$x = f(t), \qquad y = g(t), \qquad z = h(t). \tag{1}$$

The functions $f$, $g$, $h$ have some common domain of definition on the $t$-axis. In most of the cases we consider, these functions will have continuous first derivatives, and, except possibly for isolated exceptional values of $t$, the three derivatives $f'(t)$, $g'(t)$, $h'(t)$ will not all be 0 for the same value of $t$. If we think of $t$ as a variable which measures time, equation (1) describes the motion of the point $(x, y, z)$ along a path in space. We may extend the notions about vectors and velocity (see Chapter XIII) from two to three dimensions. Then

$$\mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \tag{2}$$

is the position vector of the moving point, and

$$\mathbf{V} = \frac{d\mathbf{R}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k} \tag{3}$$

is the vector velocity of the point in its path. Our assumption that $f'(t)$, $g'(t)$, $h'(t)$ are not all 0 at once is then the assumption that $\mathbf{V} \neq \mathbf{O}$.

Just as in Chapter XIII, it is true here also that $\mathbf{V}$ has the direction of the tangent to the path at $(x, y, z)$. Hence the direction of the tangent to the curve for a given $t$ is

$$f'(t):g'(t):h'(t), \tag{4}$$

provided the derivatives are not all 0. This result about the direction of the tangent is valid generally, *regardless of the interpretation placed on t.* That is, if a curve is defined by (1), where $t$ is any parameter, the direction of the tangent to the curve is given by (4) provided that $f'$, $g'$, $h'$ are continuous and not all 0 at once.

**Example 1:** The curve

$$x = a\cos\theta, \quad y = a\sin\theta, \quad z = b\theta, \tag{5}$$

with $a$ and $b$ positive and $\theta$ the parameter, is called a *cylindrical helix.*

From (5) we see that $x^2 + y^2 = a^2$; this means that the curve lies on the cylinder $x^2 + y^2 = a^2$. If we imagine that $\theta$ varies with time by the formula $\theta = \omega t$, where $\omega$ is a positive constant, then $(x, y, z)$ moves around the cylinder with constant angular velocity $\omega$, and along the cylinder with constant linear speed

$$\frac{dz}{dt} = b\frac{d\theta}{dt} = b\omega.$$

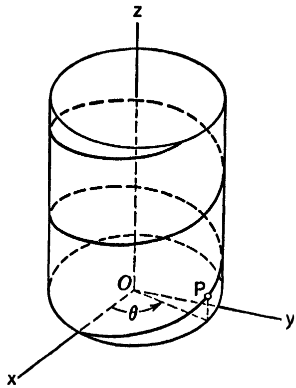



**Fig. 18-18**

The general character of the curve and the location of the point $P(x, y, z)$ in relation to the geometric representation of $\theta$ are shown in Fig. 18-18. The direction of the tangent at $P$ is

$$\frac{dx}{d\theta}:\frac{dy}{d\theta}:\frac{dz}{d\theta} = -a\sin\theta : a\cos\theta : b. \tag{6}$$

*A Curve as the Intersection of Surfaces*

It may happen that two surfaces intersect in such a way that the points of intersection form one or more curves. The equations of the two surfaces, taken simultaneously, define the locus of all points which are on both surfaces. If we confine attention to the part of this locus sufficiently near one of its points, it is often possible to regard this part of the locus as a curve defined parametrically, with some one of the coordinates as parameter.

**Example 2: In the case of Fig. 18-11, the first octant part of the intersection of the two cylinders $x^2 + z = 4$, $y = x^2$, from (0, 0, 4) to (2, 4, 0), can**

be parametrized as follows, with $x$ as the parameter:

$$x = x, \quad y = x^2, \quad z = 4 - x^2, \quad 0 \le x \le 2. \tag{7}$$

If $y$ is chosen as parameter, this same arc is represented as follows:

$$x = \sqrt{y}, \quad y = y, \quad z = 4 - y, \quad 0 \le y \le 4. \tag{8}$$

**Example 3:** Consider the intersection of the sphere $x^2 + y^2 + z^2 = 4a^2$ and the cylinder $(x - a)^2 + y^2 = a^2$.

See Fig. 18-19, in which one-fourth of the intersection is shown. The total curve is rather like a figure 8 bent so as to fit on the sphere. The curve crosses

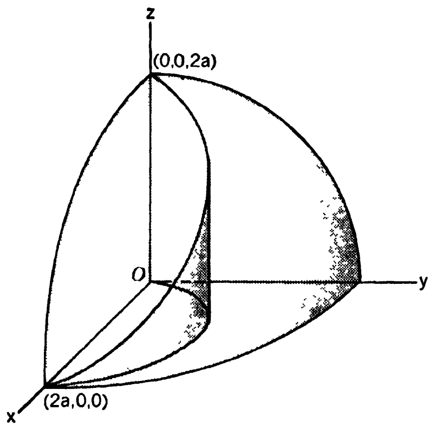



**Fig. 18-19**

itself at $(2a, 0, 0)$. For the part shown in Fig. 18-19 either $x$ or $z$ can be taken as the parameter. But $y$ cannot be used as a parameter for the whole of the part shown, because there are two different points with the same $y$-coordinate if $0 \le y < a$.

Without committing ourselves to any one choice of parameter, we can say that the direction of the tangent to the curve is

$$dx : dy : dz \tag{9}$$

unless all three differentials are zero at once. From the given equations we see that

$$2x\,dx + 2y\,dy + 2z\,dz = 0, \qquad 2(x - a)\,dx + 2y\,dy = 0.$$

Hence

$$\left.\begin{aligned} x\,dx + y\,dy + z\,dz &= 0, \\ (x - a)\,dx + y\,dy \qquad &= 0. \end{aligned}\right\} \tag{10}$$

These two homogeneous linear equations in $dx$, $dy$, $dz$ determine the direction

$$dx : dy : dz = \begin{vmatrix} y & z \\ y & 0 \end{vmatrix} : \begin{vmatrix} z & x \\ 0 & x - a \end{vmatrix} : \begin{vmatrix} x & y \\ x - a & y \end{vmatrix},$$

by the scheme of Theorem 17-A. Hence

$$dx{:}dy{:}dz = -yz{:}z(x - a){:}ay, \tag{11}$$

provided that the three quantities are not all zero. For instance, at (0, 0, 2$a$) we have

$$dx{:}dy{:}dz = 0{:}-2a^2{:}0 = 0{:}1{:}0.$$

This means that at (0, 0, 2$a$) the tangent is parallel to the $y$-axis. The ratios in (11) do not apply at (2$a$, 0, 0), however. For a discussion of the direction of the tangent at this point see Exercise 10(c).

### *The Length of a Curve*

The discussion of arc length in § 11-1 extends readily to curves in space. The basic result is that, with the parametric representation discussed at the beginning of this section, the arc length $L$ from $t_0$ to $t_1$ is

$$L = \int_{t_0}^{t_1} \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]^{1/2} dt. \tag{12}$$

If $s$ is arc length as a function of $t$, measured from some selected point on the curve and in a specified sense along the curve, then

$$ds^2 = dx^2 + dy^2 + dz^2. \tag{13}$$

If $t$ is interpreted as time, the magnitude of the vector velocity **V** is $\left|\frac{ds}{dt}\right|$; and $\frac{ds}{dt}$ itself may be either positive or negative, depending upon which way the point is moving in relation to the sense of increasing $s$ along the curve.

**Example 4:** Find the length of the curve described in (7), Example 2. With $x$ as parameter,

$$ds^2 = dx^2 + (2x\,dx)^2 + (-2x\,dx)^2 = (1 + 8x^2)\,dx^2,$$

and hence
$$L = \int_0^2 \sqrt{1 + 8x^2}\,dx.$$

This works out to be

$$L = \sqrt{33} + \frac{\sqrt{2}}{8}\log\,(4\sqrt{2} + \sqrt{33}).$$

## EXERCISES

**1.** Consider the curve $x = 6t$, $y = 3\sqrt{2}\,t^2$, $z = 2t^3$. (a) Find the direction of its tangent at $t = 1$, and find where this tangent pierces the plane $z = 0$. (b) Find the length of the arc of the curve from $t = 0$ to $t = 2$.

**2.** Consider the curve $x = 3t - t^3$, $y = 3t^2$, $z = 3t + t^3$. (a) Find the direction cosines and direction angles of the velocity vector at $t = 1$ (interpreting $t$ as time). (b) Find the equation of the plane perpendicular to the curve at $(-2, 12, 14)$. (c) Find the length of the curve from $t = 0$ to $t = 3$.

**3.** Let $C$ be the curve of intersection of the parabolic cylinders $4y = x^2$, $12z = x^2$. (a) Show that $C$ lies in a plane, and find the direction of the normal to this plane. (b) Find the direction of the tangent to $C$ at $t = 1$. (c) If a point $P$ is moving on the curve in such a way that its $z$-coordinate is increasing 2 units per second, find the vector velocity of $P$ at $x = -2$.

**4.** Consider the intersection of the surfaces $x^2 + (y - 8)^2 = 64$, $9x^2 + 16x^2 = 144y$. (a) Show that the intersection is composed of two ellipses. Find the planes in which they lie, and the lengths of their major and minor axes. (b) Find the parametric representation of the part of the intersection in the first octant, with $y$ as parameter. Find the direction of the tangent to this part at $y = 12$.

**5.** (a) Show that the intersection of $9x^2 + 25y^2 = 150y$ and $800y = 48x^2 + 75z^2$ consists of two circles. Find their radii and the planes in which they lie. (b) Find the direction of the tangent to one of the circles by the method used in connection with (10) in the text. Evaluate at $(5, 3, 4)$. (c) Express the part of the intersection in the first octant parametrically, with $z$ as parameter.

**6.** (a) For the helix of Example 1, find the relation between $ds$ and $d\theta$, assuming that $s$ increases as $\theta$ increases. Then find the length of one complete turn of the helix around the cylinder. (b) Find the acute angle which the tangent to the helix at $P$ makes with the plane through $P$ parallel to the plane $z = 0$.

**7.** Consider the curve $x = a\theta \cos \theta$, $y = a\theta \sin \theta$, $z = b\theta$, where $a > 0$, $b > 0$. (a) Show that it lies on a right circular cone. (b) If $\theta$ increases at the constant rate $\omega$, find the velocity vector at $\theta = \pi/2$ and at $\theta = \pi$. (c) If $\phi$ is the acute angle between the line $OP$ produced and the tangent to the curve at the point $P$ located by $\theta$, show that $\phi$ approaches $\pi/2$ as $\theta \to \infty$.

**8.** (a) Express the curve of (7) or (8) with $z$ as parameter. (b) Calculate $ds^2$ in terms of $y$ and $dy$, and also in terms of $z$ and $dz$, for this curve.

**9.** (a) Find the length of the curve $x = a(\theta - \sin \theta)$, $y = a(1 - \cos \theta)$, $z = 4a \sin (\theta/2)$, from $\theta = 0$ to $\theta = 2\pi$. (b) Find the length of the curve $x = a \cosh t$, $y = a \sinh t$, $z = at$ from $t = 0$ to $t = 1$.

**10.** (a) For the curve of intersection shown in Fig. 18-19, at what point on the curve is the tangent perpendicular to the $y$-axis? (b) Express the curve parametrically with $x$ as the parameter. (c) Express the curve parametrically with $z$ as the parameter. (d) Using the result of (c), find the direction of the tangent at $(2a, 0, 0)$ and interpret the result. (e) Find the $y$ and $z$ coordinates of the point on the curve where the tangent makes a 60° angle with the positive $z$-axis. (f) If the curve is parametrized with $x = a(1 + \cos \theta)$, $y = a \sin \theta$, $0 \leq \theta \leq \pi$, show that $z = 2a \sin (\theta/2)$. Then express $ds^2$ in terms of $\theta$ and $d\theta$.

## Review Questions and Problems for Chapters XVI, XVII, XVIII

### CONCEPTS AND DEFINITIONS

1. What is the relation between approximation by differentials and the tangent to a curve $y = f(x)$ at $x_0$? Explain the sign of $\Delta y - dy$ from this standpoint.
2. Is there a connection between differential approximation and Taylor's formula? Explain.
3. Define the determinant function of order two. On how many variables does it depend?
4. On how many variables does the determinant function of order three depend? Define this function as a sum of products, explaining how the products are formed and how their signs are determined.
5. Define the minor of an entry in a determinant of third order. Explain the uses of minors in calculating the value of the determinant.
6. What is meant by a homogeneous linear equation?
7. What is meant by saying that the columns of a determinant are linearly dependent?
8. What is a vector in three-dimensional space?
9. What are direction cosines? What are direction components? What is meant by "the direction" of an unsensed line?
10. How can you recognize the direction of the normal to a plane by examining the equation of the plane?
11. What is the cross-product of two vectors: (a) by definition in geometric terms? (b) in terms of the components of the two vectors?
12. If the rows of a third-order determinant are viewed as sets of components of vectors, how can the value of the determinant be expressed in terms of the vectors? How does this lead, through a geometric interpretation, to a simple necessary and sufficient condition for the determinant to be equal to zero?

### THEORY

1. What is the geometrical basis of Newton's method of finding an approximate solution of the equation $f(x) = 0$? Deduce the formula of Newton's method for successive approximations to this solution.
2. (a) Explain the trapezoidal rule and work out the formula which expresses the rule. (b) Do the same for Simpson's rule.
3. State and prove Cramer's rule for a system of two equations in two unknowns.

4. Read Theorems 17-B through 17-E carefully, skipping the proofs. Then see if you can work out the proofs by yourself, referring when necessary to § 17-2 but not to § 17-3. Take plenty of time and try to think things through, getting the necessary relationships firmly in mind and then writing the arguments down clearly.

5. Suppose the system (1) in § 17-2 has a solution, but that its determinant $D$ is 0. Show that the numerator determinant of (1) in Theorem 17-F is 0, and likewise for two other analogous determinants.

6. What is the locus of a point $(x, y, z)$ that moves in such a way that each coordinate is a linear function of the time $t$, at least one function being nonconstant? Explain.

7. Define $\mathbf{A} \cdot \mathbf{B}$ in terms of components of the vectors. How can the scalar product be expressed without mentioning components? Justify.

8. What kind of equation is characteristic of a plane? State and prove a theorem on this subject.

## PROBLEMS

1. Find the point $(x_0, y_0)$ in the first quadrant on the curve $y = \cosh x$ at which the line joining $(x_0, y_0)$ to the origin is tangent to the curve. Draw an adequate figure and use Table II.

2. A rectangle of width $c$ and length $L$ is fitted inside a rectangle of width $a$ and length $b$, as shown in Fig. 18-20; it is assumed that $c < a < b$.

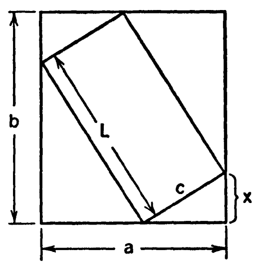


**Fig. 18-20**

(a) With the dimension $x$ as shown in the diagram, show that $x$ and $L$ satisfy the equations

$$cx + L\sqrt{c^2 - x^2} = bc, \qquad Lx + c\sqrt{c^2 - x^2} = ac,$$

and hence that $x$ is determined as the unique solution of the equation

$$\frac{2x^2 - bx - c^2}{a} = -\sqrt{c^2 - x^2}.$$

(b) Make a sketch showing how $x$ can be determined graphically from the intersection of a parabola and a circle. Do this carefully and estimate the value of $x$ if $a = 6$, $b = 14$, $c = 1$.
(c) Convert the problem into a problem of solving $f(x) = 0$, where $f(x)$ is a polynomial of degree 4. Find $x$ to four places by Newton's method if $a = 6$, $b = 14$, $c = 1$.
(d) Solve for $x$ as in (c) if $a = 10$, $b = 15$, $c = 1$.

3. Figure 18-21 shows a mechanical system in equilibrium. The lengths $OA$, $AB$, and $BC$ are equal. Equal weights are attached at $B$ and $C$ on the

string $ABC$. Point $A$ is fixed, and the weight at $C$ is constrained by a smooth ring which can slide on the vertical rod $OC$. In the equilibrium position the angles $\theta$, $\phi$ are determined by the equations

$$\tan \phi = 2 \operatorname{ctn} \theta, \qquad \sin \phi = 1 - \cos \theta.$$

If $x = \sin \phi$, show that $x$ satisfies the equation

$$3x^4 - 6x^3 + 8x - 4 = 0.$$

Find $\sin \phi$ to four decimal places.

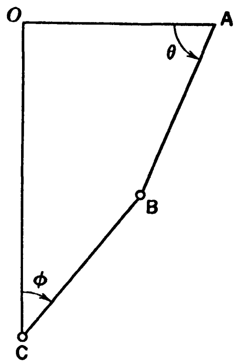



**Fig. 18-21**

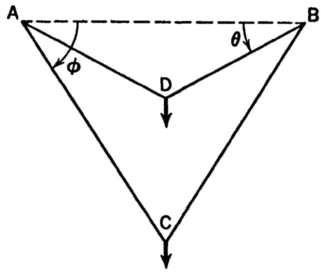



**Fig. 18-22**

4. Figure 18-22 shows a length of string 60 inches long over two smooth pegs at $A$ and $B$, 20 inches apart. The system is in equilibrium with a weight of 2 pounds at $C$ and a weight of 1 pound at $D$. The statics and geometry of the situation lead to the equations

$$\sin \phi = 2 \sin \theta, \qquad \cos \phi + \cos \theta = 3 \cos \theta \cos \phi.$$

If $x = \cos \phi$, show that $x$ is a certain root of the equation

$$3x^4 - 2x^3 + 8x^2 - 6x + 1 = 0.$$

Locate the root roughly, and then improve the approximation by Newton's method. Explain how you are sure you have the right root.

5. Find $x$ approximately if

$$\int_0^x \frac{t^2}{1 + t^2}\, dt = \frac{1}{2}.$$

With one method of solution a fairly quick answer can be obtained with the aid of a graph and Table III.

6. Find $x$ approximately if $\cos x = \sqrt{x}$. Begin with a good graph of $y = \cos x$ and $y = \sqrt{x}$, and locate the intersection roughly. Then use Table III and a table of square roots. Finally, use Newton's method to obtain a four-decimal place answer.

**7.** Show that the line in the $xy$-plane, through the distinct points $(x_1, y_1)$, $(x_2, y_2)$, has the equation

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$$

**8.** Show, by an argument like that used in getting (6) in § 18-3, that the circle in the $xy$-plane, determined by three noncollinear points $(x_1, y_1)$, $(x_2, y_2)$, $(x_3, y_3)$, has the equation

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0.$$

**9.** If $P_k$ has coordinates $(x_k, y_k)$, and if the vertices of a triangle, in counterclockwise order around the triangle, are $P_1$, $P_2$, $P_3$, show that the area of the triangle is

$$\frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

**10.** Show that

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (a - b)(b - c)(c - a).$$

**11.** Show that, if line segments are drawn joining the mid-points of opposite edges of a tetrahedron, these three line segments all pass through a single point which is the mid-point of each of them.

(a) Show that the equation of the plane through the points $(x_1, y_1, z_1)$, $(x_2, y_2, z_2)$ and perpendicular to the plane $ax + by + cz + d = 0$ is

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ a & b & c & 0 \end{vmatrix} = 0.$$

(b) What is the equation, in determinant form, of the plane through $(x_1, y_1, z_1)$ and perpendicular to each of the planes

$$a_1x + b_1y + c_1z + d_1 = 0, \qquad a_2x + b_2y + c_2z + d_2 = 0?$$

**12.** Show that

$$\begin{vmatrix} x - a & y - b & z - c \\ l & m & n \\ p & q & r \end{vmatrix} = 0$$

is the equation of a plane through $(a, b, c)$ and having its normal perpendicular to each of the directions $l{:}m{:}n$, $p{:}q{:}r$.

**13.** Derive an equation for the plane through $(a, 0, 0)$, $(0, b, 0)$, $(0, 0, c)$, where $abc \neq 0$.

**14.** If $a$, $b$, $c$ are the intercepts of a plane on the $x$, $y$, and $z$ axes, respectively, and if $p$ is the perpendicular distance of the plane from the origin, show that $1/p^2 = (1/a^2) + (1/b^2) + (1/c^2)$.

**15.** Find the equation of the plane through the point $(2, 1, -3)$ and the line $x - 3 = y + 2 = z - 1$.

**16.** Is the plane $4x + 3y - 12z = 8$ tangent to the sphere $x^2 + y^2 + z^2 - 2x + 4y + 6z + 10 = 0$?

**17.** Find the line through $(6, 4, 3)$, parallel to the plane $5x + y - 3z = 6$, and intersecting the $y$-axis.

**18.** Find the locus of all points equidistant from the $y$-axis and the plane $z = -4$.

**19.** Find the locus of all points $P$ such that the distance from $P$ to the $z$-axis is three-fifths the distance from $P$ to the origin.

**20.** Find the locus of all points $P$ such that the distance from $P$ to the plane $y = -2$ is equal to the distance from $P$ to $(0, 2, 0)$.

**21.** Find the value of the determinant

$$\begin{vmatrix} 36 & 18 & 0 & 0 & 36 \\ 18 & 6 & 0 & 0 & 24 \\ 6 & 1 & 0 & -1 & 12 \\ 0 & 0 & 18 & 6 & 24 \\ 0 & 0 & 36 & 18 & 36 \end{vmatrix}.$$

*CHAPTER* XIX

# PARTIAL DIFFERENTIATION

## 19-1 Functions of Several Variables

Functions of several variables have occurred frequently in Chapter XVIII, and they occur commonly in formulas relating to geometric figures. For example, if we have a right circular cone of altitude $h$ and radius of base $r$, its volume and lateral surface area are, respectively,

$$V = \frac{\pi}{3} r^2 h, \qquad S = \pi r \sqrt{r^2 + h^2}.$$

Here $V$ and $S$ are functions of the two variables $r$, $h$. The law of cosines expresses the length $c$ of the third side of a triangle as a function of the lengths $a$, $b$ of the other two sides and of the included angle $\theta$. The formula expressing $c$ as a function of three variables is

$$c = (a^2 + b^2 - 2ab \cos \theta)^{1/2}.$$

A function of two variables is defined as follows. Let $D$ be a collection of number pairs $(x, y)$, and suppose that with each pair $(x, y)$ is associated a unique number $z$, thus giving us a certain collection of number triples $(x, y, z)$. This collection is called a function of two variables. If we denote the function by a single letter, say $f$, then we write $z = f(x, y)$ and call $z$ the value of $f$ at $(x, y)$. We call $x$ and $y$ independent variables, and $z$ is called the dependent variable. The collection of all the values of $f$ is called the *range* of $f$, while the collection $D$ of allowable pairs $(x, y)$ is called the *domain* of $f$.

Functions of three or more variables are defined in a similar manner.

If $f$ is a function of two variables, with $z = f(x, y)$, and if we interpret

$(x, y, z)$ as rectangular coordinates of a point in space, the collection of all points obtained in this way from the function $f$ is called the *graph* of $f$. We can think of the domain of $f$ as a collection of points $(x, y)$ in the $xy$-plane. On each such point we construct the ordinate with $z = f(x, y)$, and this gives us the point of the graph.

**Example 1:** If $f(x, y) = \frac{1}{6}(12 - 3x - 4y)$, the graph is characterized by the equation

$$z = \frac{1}{6}(12 - 3x - 4y), \quad \text{or} \quad 3x + 4y + 6z = 12.$$

In this case the graph is the plane through the three points (4, 0, 0), (0, 3, 0), (0, 0, 2).

**Example 2:** If

$$f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2},$$

the graph is the elliptic paraboloid of Example 4, § 8-6 (see Fig. 18-14).

*The Definition of a Limit*

Suppose the domain of definition of $f$ includes some points $(x, y)$ as close to $(a, b)$ as one pleases, though the domain need not necessarily contain $(a, b)$ itself. In other words, we suppose that $(x, y)$ can approach $(a, b)$ while remaining in the domain of $f$. The point $(a, b)$ might be completely surrounded by the domain of $f$, or it might be at an edge or corner of the domain of $f$.

DEFINITION. We say that $f(x, y)$ approaches the number $A$ as limit when $(x, y)$ approaches $(a, b)$ if $|f(x, y) - A|$ can be made as small as we please merely by requiring that $(x, y)$ be in the domain of $f$ and sufficiently near $(a, b)$, though distinct from $(a, b)$. This limiting behavior of $f$ is expressed by writing

$$\lim_{(x,y)\to(a,b)} f(x, y) = A. \tag{1}$$

**Example 3:** Suppose

$$f(x, y) = \frac{1}{\log (x^2 + y^2)}.$$

Here the domain of $f$ consists of all $(x, y)$ such that $0 < x^2 + y^2$ and $x^2 + y^2 \neq 1$. In this case

$$\lim_{(x,y)\to(0,0)} f(x, y) = 0.$$

Suppose, for example, that $\epsilon > 0$ and that we wish to make $|f(x, y) - 0| < \epsilon$. This is equivalent to

$$\frac{1}{\epsilon} < |\log (x^2 + y^2)| = \log \frac{1}{x^2 + y^2}$$

if $0 < x^2 + y^2 < 1$. This in turn is equivalent to

$$e^{1/\epsilon} < \frac{1}{x^2 + y^2}, \quad \text{or} \quad 0 < x^2 + y^2 < e^{-1/\epsilon}.$$

Thus, to make $|f(x, y) - 0| < \epsilon$, it is sufficient to have $(x, y)$ in the domain of $f$ and at a distance less than $e^{-1/2\epsilon}$ from the origin.

*Continuity*

Let $f$ be a function of $x$ and $y$ and let $(a, b)$ be a point of its domain. Then $f$ is called continuous at $(a, b)$ if $f(x, y)$ approaches the limit $f(a, b)$ when $(x, y)$ approaches $(a, b)$.

Our main concern in this chapter is with functions which are continuous. Points of discontinuity are therefore exceptional, so far as we are at present concerned. Some exceptions do occur.

**Example 4:** Let a function $f$ be defined as follows. Its domain consists of all $(x, y)$ except $(0, 0)$. For a given $(x, y)$ in this domain the value $f(x, y)$ is defined to be the radian measure $\theta$ of the angle such that $0 \leq \theta < 2\pi$ and

$$\sin \theta = \frac{y}{\sqrt{x^2 + y^2}}, \qquad \cos \theta = \frac{x}{\sqrt{x^2 + y^2}}.$$

That is, $\theta$ is the polar coordinate angle for the point $(x, y)$, chosen with the restriction that $0 \leq \theta < 2\pi$.

It is intuitively clear that $\theta$ is a continuous function of $(x, y)$ at $(a, b)$ if $(a, b)$ is not the origin or on the positive $x$-axis. But, if $a > 0$ and $b = 0$, $f$ is discontinuous at $(a, b)$. For then $f(a, b) = 0$ but $f(x, y)$ does not approach 0 as $(x, y) \rightarrow (a, b)$. The student should see, for instance, that $f(x, y)$ does not approach any limit as $(x, y)$ approaches $(1, 0)$. At some points near $(1, 0)$ the value of $f(x, y)$ is 0 or near 0; but at other points near $(1, 0)$ the value of $f(x, y)$ is near $2\pi$.

If two functions $f$ and $g$ are both continuous at $(a, b)$, so are the sum and product functions $f(x, y) + g(x, y)$, $f(x, y) \cdot g(x, y)$, and so also is the quotient function $f(x, y)/g(x, y)$, provided that $g(a, b) \neq 0$. Functions which are constructed by composition of continuous functions are again continuous, under appropriate specifications. For example, $\sin u$ is a continuous function of $u$ at $u = 0$; $x^2 + y^2$ is a continuous function of $(x, y)$ at $(0, 0)$, with value 0 there. Hence $\sin (x^2 + y^2)$ is continuous at $(0, 0)$.

Limits and continuity are defined in much the same way in the case of functions of three or more variables.

*Level Curves*

If $f$ is a function of $x$ and $y$, and $k$ is a constant, it frequently happens in practice that the locus of all points $(x, y)$ such that $f(x, y) = k$ is a curve in the $xy$-plane. It is called a *level curve* of the function. If we know the level curves for various values of $k$, we can obtain a very good idea of what the function is like. The representation of a function by drawing level curves of it is based on the same idea as that which is used in representing the configuration of the land surface in a certain region by a topographical map of the region.

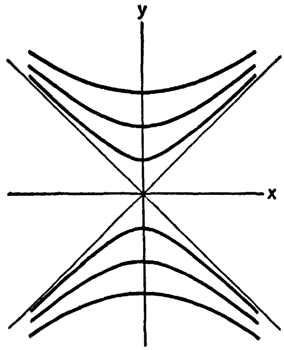



**Fig. 19-1**

**Example 5:** If $f(x, y) = \sqrt{y^2 - x^2}$, the level curves are defined by $\sqrt{y^2 - x^2} = k$.

The only admissible values of $k$ are positive or zero. For $k = 0$ we get $y^2 - x^2 = 0$, which represents the two lines $y = \pm x$. For $k > 0$ the level curves are rectangular hyperbolas as shown in Fig. 19-1. The curves shown correspond to $k = 1, 2, 3$. The graph of $z = \sqrt{y^2 - x^2}$ is the portion of the conical surface $z^2 = y^2 - x^2$, or $x^2 + z^2 = y^2$, on which $z \geq 0$. It is a right circular cone with axis along the $y$-axis. The level curve $\sqrt{y^2 - x^2} = k$ is just like the curve in which the plane $z = k$ intersects the cone.

In the general case, the level curves $f(x, y) = k$ are just like the curves in which the various planes $z = k$ intersect the surface defined by $z = f(x, y)$.

*Level Surfaces*

If $f$ is a function of three variables $x, y, z$, and if we write $w = f(x, y, z)$, a graphical representation of the function can be made by talking about points $(x, y, z, w)$ in space of four dimensions. But physical intuition about functions of three variables may be better served by using the notion of a *level surface*. For a given constant $k$ the locus of points $(x, y, z)$ in three-dimensional space such that $f(x, y, z) = k$ may be a surface. If so, we call it a *level surface*. By visualizing the various level surfaces, we can form an idea of the nature of the function.

**Example 6:** Let $f(x, y, z) = \dfrac{x^2}{25} + \dfrac{y^2}{16} + \dfrac{z^2}{9}$.

Here the admissible values of $k$ are those for which $k \geq 0$. If $k = 0$, there is no level surface; the locus $f(x, y, z) = 0$ is the single point $(0, 0, 0)$. If $k > 0$, the level surface is an ellipsoid. All these ellipsoids have the same center and the same axes of symmetry. As we go out away from the origin, the values of $f$ increase. As we shall see later, the direction of most rapid increase at a point is the direction perpendicular at that point to the ellipsoid which is the level surface.

## 19-2 Partial Derivatives

If $F$ is a function of three variables $x$, $y$, $z$, we can obtain a function of one variable by assigning fixed values to the other two variables. If $y$ and $z$ are regarded as fixed, the derivative of $F$ with respect to $x$ is called the *partial derivative* of $F$ with respect to $x$. This partial derivative is denoted by $\partial F/\partial x$. To indicate its value at $x = a$, $y = b$, $z = c$, we can use the symbol

$$\left(\frac{\partial F}{\partial x}\right)_{(a,b,c)}.$$

The symbol $\partial F/\partial x$ alone is usually understood to denote either the value of the partial derivative at $(x, y, z)$ or the partial derivative as a function of $x$, $y$, $z$. Similar notations are used for the partial derivatives with respect to $y$ and $z$. This same kind of notation is used, regardless of how many independent variables there are.

**Example 1:** If

$$F(x, y, z) = \log (x^2 + y^2) + \sqrt{y^2 + z^2} - \frac{xy}{z},$$

then

$$\frac{\partial F}{\partial x} = \frac{2x}{x^2 + y^2} - \frac{y}{z}, \qquad \frac{\partial F}{\partial z} = \frac{z}{\sqrt{y^2 + z^2}} + \frac{xy}{z^2}.$$

Let us consider a function $f$ of two independent variables $x$, $y$. We denote the dependent variable by $z$ and consider the graph of $z = f(x, y)$, on the assumption that $f$ is continuous at each point $(x, y)$ in some rectangular neighborhood of the point $(a, b)$. By a "rectangular neighborhood" of $(a, b)$ we mean the part of the $xy$-plane inside some rectangle having its center at $(a, b)$ and each of its sides parallel to a coordinate axis.

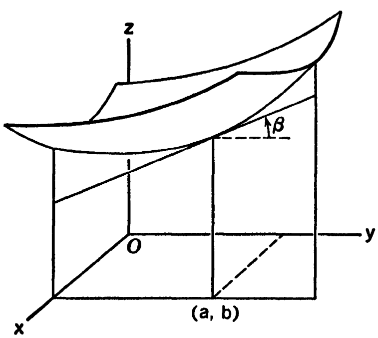


**Fig. 19-2**

We propose to show the geometrical significance of the value of $\partial f/\partial y$ at $(a, b)$. Since $z = f(x, y)$, we often write $\partial z/\partial y$ in place of $\partial f/\partial y$. If we keep $x$ constant, say $x = a$, and regard $y$ as a variable, then the points $(a, y, z)$ given by $z = f(a, y)$ form the curve in which the plane $x = a$ intersects the surface $z = f(x, y)$. If this curve has a tangent at the point where $y = b$, the value of $\partial z/\partial y$ at this point is $\tan \beta$, where $\beta$ is the angle from the positive $y$-direction to the tangent; see Fig. 19-2. There is, of course, a similar geometrical interpretation of the value of $\partial z/\partial x$ at $(a, b)$.

*The Tangent Plane at a Point of $z = f(x, y)$*

Let $c = f(a, b)$. We are going to suppose that the surface $z = f(x, y)$ has a tangent plane at $(a, b, c)$, and that this plane is not parallel to the $z$-axis. From this we shall show that $f$ has partial derivatives with respect to $x$ and $y$, respectively, when $x = a$ and $y = b$, and we shall show how to find the equation of the tangent plane.

First of all, we must know what it means for a plane through $(a, b, c)$ to be tangent to the surface there. Let $(x, y, z)$ be any point other than $(a, b, c)$ of the surface, and let $L$ be the line through $(a, b, c)$ and $(x, y, z)$. Let $M$ be a fixed plane through $(a, b, c)$ and let $\theta$ be the angle (not over $\pi/2$) between $L$ and $M$ (see Fig. 19-3). Then $M$ is tangent to the surface at $(a, b, c)$ if $\theta$ approaches 0 as $(x, y, z)$ approaches $(a, b, c)$ on the surface.

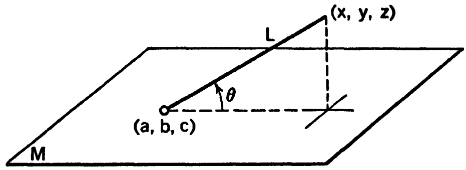



**Fig. 19-3**

Now, if $M$ is the plane tangent to $z = f(x, y)$ at $(a, b, c)$, and if $M$ is not parallel to the $z$-axis, it is evident from the definition that the plane $x = a$ must intersect $M$ in a straight line which is tangent at $(a, b, c)$ to the curve $x = a$, $z = f(a, y)$ (the curve and tangent illustrated in Fig. 19-2). The equation of $M$ can be written in the form

$$z - c = A(x - a) + B(y - b),$$

where $A$ and $B$ are certain constants. The line of intersection with $x = a$ has the equations

$$x = a, \qquad z - c = B(y - b).$$

Hence $B$ must be the value of $\partial f/\partial y$ when $x = a$ and $y = b$, since it is clear that $B$ must be equal to $\tan \beta$, where $\beta$ is the angle in Fig. 19-2. For exactly similar reasons, the value of $\partial f/\partial x$ at this point must be $A$. Therefore the equation of the tangent plane is

$$z - f(a, b) = \left(\frac{\partial f}{\partial x}\right)_{(a,b)}(x - a) + \left(\frac{\partial f}{\partial y}\right)_{(a,b)}(y - b). \qquad (1)$$

We see then that if the surface $z = f(x, y)$ has a tangent plane not parallel to the $z$-axis at $(x, y, z)$, the line normal to this plane has the direction

$$\frac{\partial f}{\partial x} : \frac{\partial f}{\partial y} : -1. \qquad (2)$$

The line through $(x, y, z)$ with this direction is called the normal to the surface at this point.

**Example 2:** Find the plane tangent to the paraboloid

$$z - 3 = -\left(\frac{x^2}{9} + \frac{y^2}{16}\right)$$

at $x = 2$, $y = 2$, and the direction of the normal to the paraboloid at this point.

In this case

$$\frac{\partial z}{\partial x} = -\frac{2x}{9} = -\frac{4}{9}, \qquad \frac{\partial z}{\partial y} = \frac{-2y}{16} = -\frac{1}{4}$$

at the point in question. The value of $z$ is found to be 83/36. Hence the equation of the tangent plane is

$$z - \frac{83}{36} = -\frac{4}{9}(x - 2) - \frac{1}{4}(y - 2),$$

or

$$16x + 9y + 36z = 133.$$

The direction of the normal is

$$-\frac{4}{9}:-\frac{1}{4}:-1, \quad \text{or} \quad 16:9:36.$$

If the equation of the surface defines $z$ implicitly, rather than explicitly, as a function of $x$ and $y$, we calculate the partial derivatives of $z$ by the method of implicit functions as it was described for functions of one variable in § 3-7.

**Example 3:** Find the equation of the plane tangent to the surface $z^3 + 3xz - 2y = 0$ at $(1, 7, 2)$.

To find $\dfrac{\partial z}{\partial x}$ we have

$$3z^2 \frac{\partial z}{\partial x} + 3x \frac{\partial z}{\partial x} + 3z = 0, \qquad \frac{\partial z}{\partial x} = -\frac{z}{z^2 + x}.$$

Likewise,

$$3z^2 \frac{\partial z}{\partial y} + 3x \frac{\partial z}{\partial y} - 2 = 0, \qquad \frac{\partial z}{\partial y} = \frac{2}{3(z^2 + x)}.$$

Evaluating at $(1, 7, 2)$, we have

$$\frac{\partial z}{\partial x} = -\frac{2}{5}, \qquad \frac{\partial z}{\partial y} = \frac{2}{15}.$$

The tangent plane is

$$z - 2 = -\frac{2}{5}(x - 1) + \frac{2}{15}(y - 7),$$

or

$$6x - 2y + 15z = 22.$$

## EXERCISES

**1.** Compute all the first partial derivatives of each function.

(a) $f(x, y) = \sqrt{xy} + \sin xy^3 + \cos x^2y.$

(b) $f(x, y) = e^{\sin xy} + e^{\cos(x+y)}.$

(c) $f(x, y) = (x^3 - 2y)^2 + \sqrt{x^2 - xy}$.

(d) $f(x, y) = x^2 \tan^{-1} \dfrac{y}{x}$.

(e) $F(x, y, z) = \tan\left(\dfrac{x}{y} + \dfrac{y}{z} + \dfrac{z}{x}\right)$.

(f) $G(r, \theta, \varphi) = r^2 \sin\theta \cos\varphi + \dfrac{\sin\varphi}{r}$.

(g) $F(a, b, \theta) = (a^2 + b^2 - 2ab\cos\theta)^{1/2}$.

**2.** If $z$ is defined implicitly by the given equation, find $\partial z/\partial x$ and $\partial z/\partial y$ at the point indicated.
(a) $4x^2 + 2xz + z^2 - yz = 1$ at $(1, -2, -1)$.
(b) $x^2y + xz^2 + y^3 + z^3 = 28$ at $(3, 2, -1)$.
(c) $4(x^2 + y^2) - (z - 5)^2 = 0$ at $(3, 4, -5)$.
(d) $(x^2 + y^2 + z^2)^3 = 216z^2$ at $(1, 2, 1)$.
(e) $x^2 \cos^2 z - y^2 \sin^2 z = \sin^2 2z$ at $(1, 0, \pi/6)$.
(f) $z = \log(x^2 + yz + z^2 - 4)$ at $(e, 0, 2)$.

**3.** Find $\dfrac{\partial F}{\partial x} + \dfrac{\partial F}{\partial y} + \dfrac{\partial F}{\partial z}$ if $F(x, y, z) = xy^2 + yz^2 + zx^2$.

**4.** Find $x \dfrac{\partial f}{\partial x} + y \dfrac{\partial f}{\partial y}$ if $f(x, y) = y/x$.

**5.** If $z = y^2 + \tan(ye^{1/x})$, show that $x^2 \dfrac{\partial z}{\partial x} + y \dfrac{\partial z}{\partial y} = 2y^2$.

**6.** Find $2x \dfrac{\partial F}{\partial x} + y \dfrac{\partial F}{\partial y} + 2z \dfrac{\partial F}{\partial z}$ if $F(x, y, z) = x \sin(y^2/z) - y^2 \tan(z/y^2)$.

**7.** Find where the tangent to the curve $x = 1$, $x^2 + y^2 - 2y + 4z + 8 = 0$ is parallel to the $xy$-plane.

**8.** Find (a) the angle which the line tangent to the curve $x = 3$, $18z = 4x^2 + 9y^2$ at $(3, 2, 4)$ makes with the $xy$-plane; (b) the angle which the line tangent to the curve $y = 2$, $18z = 4x^2 + 9y^2$ at $(3, 2, 4)$ makes with the $xy$-plane; (c) the angle which the line tangent to the curve $z = 3$, $x^3 - 2x^2z^2 + 2xz^3 + 2yz = 0$ at $(2, -\frac{22}{3}, 3)$ makes with the $xz$-plane.

**9.** Find the equation of the tangent plane, and the direction of its normal, in the case of each surface at the point indicated.

(a) $z = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$ at $(a, b, 2)$.

(b) $(x^2/16) + z = y^2/9$ at $(15, \frac{15}{4}, 25)$.
(c) $5x^2 + 4y^2 + 2z^2 = 17$ at $(-1, 1, 2)$.
(d) $x^2 + y^2 + z^2 - 4z = 10$ at $(1, 2, -1)$.
(e) $z^3 + 3xz - 2y = 0$ at $(1, 7, 2)$.
(f) $z^3 + y^3 + z^3 - 3xyz = 8$ at $(3, 3, 2)$.

10. In growing a certain agricultural crop, it is found that within certain limits the yield $z$ in bushels per acre is given by the formula $z = 50(3x - 2x^2/y)$, where $x$ is the number of plants grown per square foot, and $100y$ is the number of man-hours expended in caring for the crop. (a) Draw the three level curves for $z$ as a function of $x$ and $y$, through the points (1, 1), (2, 2), and (3, 3), respectively. (b) For a yield of 150 bushels per acre, how many plants per square foot are needed to require the smallest amount of care? (c) If 400 man-hours of labor are available, how many plants per square foot should be drawn to insure a maximum yield?

11. In growing a certain crop the total yield from $y$ acres is found to be $z$ tons, where $z = \dfrac{24xy - 5x^2 - 16y^2}{x + y}$ and $100x$ is the number of man-hours of labor that are employed. (a) For a fixed plot of 5 acres, how many man-hours of labor will produce a maximum yield? (b) Suppose 800 man-hours of labor and 4 acres are devoted to the crop. Describe quantitatively the effects of a change in labor time with fixed acreage, and of a small increase in acreage with fixed labor time. (c) Find the slope of the level curve through the point $x = 8$, $y = 5$ in the $xy$-plane. What meaning does this slope have with reference to maintaining a constant yield, if small changes are made in $x$, near $x = 8$?

12. (a) Find the tangent plane to the hyperbolic paraboloid $144z = 9y^2 - 16x^2$ at $(4, \frac{20}{3}, 1)$. (b) Show that, if $(x, y, z)$ is on the line of intersection of the planes $4x + 3y = 36$, $4x - 3y = -4z$, it is on the hyperbolic paraboloid. (c) Show that, if $(x, y, z)$ is on the line of intersection of the planes $4x - 3y = -4$, $4x + 3y = 36z$, it is on the hyperbolic paraboloid. (d) Verify directly by analytic geometry that the plane determined by the lines in (b) and (c) is the same as the tangent plane in (a).

## 19-3 The Differential of a Function of Several Variables

In the case of a function of one variable we recall from § 5-1 that the differential was defined as follows. Suppose that we are dealing with $y = f(x)$ at $x = a$, where $f$ is assumed to have a derivative. Then we take $dx$ to be an independent variable and we define $dy$ as a linear function of $dx$ by the formula $dy = f'(a)\,dx$. This linear function of $dx$ is called the differential of $f$ at $x = a$.

In the two-variable case of $z = f(x, y)$ we generalize the differential concept with the following idea in mind. For the differential of $f$ at $x = a$, $y = b$, we want $dz$ to be a linear function of $dx$ and $dy$ of the type $dz = A\,dx + B\,dy$, where $A$ and $B$ are certain numbers. Now, in the one-variable case, $dy$ and $dx$ were related in such a way that $dy/dx = f'(a)$ when $dx \neq 0$. Here then it is natural to want

$$\frac{dz}{dx} = \left(\frac{\partial f}{\partial x}\right)_{(a,b)}$$

if $dx \neq 0$ and $dy = 0$, with a symmetrical requirement if $dy \neq 0$ and $dx = 0$. This means we must choose

$$A = \left(\frac{\partial f}{\partial x}\right)_{(a,b)}, \qquad B = \left(\frac{\partial f}{\partial y}\right)_{(a,b)}. \tag{1}$$

One might be inclined to suppose that this is as far as we need to go in defining a differential. That is, one might think we could suppose merely that $f$ has partial derivatives with respect to $x$ and $y$ at $(a, b)$, and then define the differential as the linear function of $dx$ and $dy$ given by

$$dz = A\,dx + B\,dy, \tag{2}$$

where $A$ and $B$ are given by (1). But it turns out on deeper investigation that it is not satisfactory to assume so little about $f$. In order for the differential to have useful properties it has been found by experience to be wiser to make a stronger assumption.

In order to understand the reason for the stronger assumption, let us go back again to the one-variable case. If $y = f(x)$, one thing of great importance about $dy$ is that if $\Delta y = f(a + dx) - f(a)$, then $dy$ is a "good approximation" to $\Delta y$ when $dx$ is small, in the following sense: $\Delta y - dy$ is small *in comparison with* $dx$ when $dx$ is small; in fact,

$$\lim_{dx \to 0} \frac{\Delta y - dy}{dx} = 0. \tag{3}$$

This is true because

$$\frac{\Delta y - dy}{dx} = \frac{f(a + dx) - f(a)}{dx} - f'(a)$$

and

$$\lim_{dx \to 0} \frac{f(a + dx) - f(a)}{dx} = f'(a).$$

For the two-variable case we want a suitable analogue of (3). That is, if

$$\Delta z = f(a + dx, b + dy) - f(a, b), \tag{4}$$

we want $dz$ to be a "good approximation" to $\Delta z$ in a suitable sense. What we want is that $\Delta z - dz$ shall be small *in comparison with both* $dx$ *and* $dy$ when these are small. Now, one way to make both $dx$ and $dy$ small is to make $|dx| + |dy|$ small. Since

$$\sqrt{dx^2 + dy^2} \leq |dx| + |dy| \leq \sqrt{2}\sqrt{dx^2 + dy^2}, \tag{5}$$

as is easily seen (see Exercise 13), it would be equivalent if we made $\sqrt{dx^2 + dy^2}$ small. Hence the condition we desire is that

$$\lim_{|dx| + |dy| \to 0} \frac{\Delta z - dz}{|dx| + |dy|} = 0. \tag{6}$$

DEFINITION. If $\Delta z$ is defined by (4), if $dz$ is defined by (2) and (1), and if condition (6) is satisfied, we shall say that $f$ is *differentiable* at $(a, b)$. Then $dz$, as a linear function of $dx$ and $dy$, is called the differential.

Condition (6) is the stronger condition referred to earlier. There are functions for which this condition is not satisfied, even though the partial derivatives $\partial f/\partial x$, $\partial f/\partial y$ are defined at $(a, b)$. For an example, see Exercise 12. However, the following theorem can be proved.

THEOREM 19-A. *Suppose $f$, $\partial f/\partial x$, and $\partial f/\partial y$ are defined at $(a, b)$ and at all nearby points. Suppose also that the partial derivatives are continuous at $(a, b)$. Then the function $f$ is differentiable at $(a, b)$.*

It can also be shown that the geometrical meaning of $f$ being differentiable is exactly this: that the surface $z = f(x, y)$ has at the point $x = a$, $y = b$ a tangent plane not parallel to the $z$-axis. For a proof of Theorem 19-A and a fuller discussion of the subject of differentiability, see the early part of Chapter 7 in the author's text, *Advanced Calculus* (Ginn & Company, Boston, 1955).

**Example 1:** If $z = xy$, verify that condition (6) is satisfied for arbitrary choice of $(a, b)$.

Here

$$\frac{\partial z}{\partial x} = y, \qquad \frac{\partial z}{\partial y} = x.$$

At $x = a$, $y = b$ we have $dz = b\,dx + a\,dy$ and

$$\Delta z = (a + dx)(b + dy) - ab = b\,dx + a\,dy + dx\,dy.$$

Hence we have to show that

$$\lim_{|dx| + |dy| \to 0} \frac{dx\,dy}{|dx| + |dy|} = 0. \tag{7}$$

Now it is certainly true that

$$|dx||dy| < (|dx| + |dy|)^2 \tag{8}$$

if $dx$ and $dy$ are not both zero. For, if we expand the right member in (8), we see that (8) is equivalent to

$$0 < |dx|^2 + |dx||dy| + |dy|^2,$$

which is certainly true. From (8) we have

$$\frac{|dx\,dy|}{|dx| + |dy|} < |dx| + |dy|,$$

and from this it is clear that (7) is true.

The discussion of differentials can be extended in a natural way to functions of more than two independent variables. Thus if $u = F(x, y, z)$, we say that $F$ is differentiable at $(x, y, z)$ if

$$\frac{\Delta u - du}{|dx| + |dy| + |dz|} \tag{9}$$

approaches 0 as $|dx| + |dy| + |dz|$ approaches 0, where

$$\Delta u = F(x + dx, y + dy, z + dz) - F(x, y, z)$$

and

$$du = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy + \frac{\partial F}{\partial z} dz.$$

The condition for differentiability can be expressed as follows: Let $\epsilon$ be defined as the ratio in (9) if $|dx| + |dy| + |dz| \neq 0$, and let $\epsilon = 0$ if $dx = dy = dz = 0$. Then we can write

$$\Delta u = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy + \frac{\partial F}{\partial z} dz + \epsilon(|dx| + |dy| + |dz|), \tag{10}$$

and $\epsilon$ is a function of the variables $dx$, $dy$, $dz$ which is continuous at $dx = 0$, $dy = 0$, $dz = 0$.

The following general rules are valid if $u$ and $v$ are differentiable functions of the same set of independent variables:

$$dc = 0,\ c \text{ a constant}, \qquad d(u + v) = du + dv,$$

$$d(cu) = c\,du, \qquad d(uv) = u\,dv + v\,du,$$

$$d\left(\frac{u}{v}\right) = \frac{v\,du - u\,dv}{v^2}.$$

Differentiation formulas such as

$$d(u^n) = nu^{n-1}\,du, \qquad d \log u = \frac{du}{u},$$

remain true even when $u$ is a function of several independent variables. For example, if $u = f(x, y)$ and $v = u^n$, then

$$dv = \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy$$

and

$$\frac{\partial v}{\partial x} = nu^{n-1} \frac{\partial u}{\partial x}, \qquad \frac{\partial v}{\partial y} = nu^{n-1} \frac{\partial u}{\partial y},$$

whence

$$dv = d(u^n) = nu^{n-1}\left(\frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy\right) = nu^{n-1}\,du.$$

The general technique of estimation of small errors by differentials is similar to the technique illustrated in § 16-1 in the case of functions of one variable.

**Example 2:** If $u$ is computed from the formula $u = x^2y^3z^{-4}$, where $x$, $y$, $z$ are assigned positive values, by approximately what percentage might $u$ be changed if $x$, $y$, $z$ were changed by 1.5%, 1%, and 0.5%, respectively?

We use logarithms:

$$\log u = 2 \log x + 3 \log y - 4 \log z.$$

Then $$\frac{du}{u} = 2\frac{dx}{x} + 3\frac{dy}{y} - 4\frac{dz}{z}.$$

No information is given as to whether the changes in $x$, $y$, $z$ would be increases or decreases. All we are told is that

$$\left|\frac{dx}{x}\right| \le 0.015, \quad \left|\frac{dy}{y}\right| \le 0.01, \quad \left|\frac{dz}{z}\right| \le 0.005.$$

Hence, in the most unfavorable case, if all the changes worked to enlarge the change in $u$ instead of partially offsetting each other, we would have

$$\left|\frac{du}{u}\right| \le 0.03 + 0.03 + 0.02 = 0.08.$$

Thus the change in $u$ might be as great as approximately 8%.

## EXERCISES

1. Compute $\Delta u - du$ if $u = xyz$, at $x = a$, $y = b$, $z = c$.

2. Compute $\Delta z - dz$ in each case, with $dx$ and $dy$ arbitrary, and $x$, $y$ as specified. Notice that, after simplification, the answer involves $dx$ and $dy$ in such a way, through terms of degree 2 or 3, that condition (6) is satisfied.
   (a) $z = x^3 + y^3 + 3xy$, $x = 1$, $y = -1$.
   (b) $z = \dfrac{x - y}{x + y}$, $x = 8$, $y = 12$.
   (c) $z = \dfrac{xy}{x - y}$, $x$, $y$ arbitrary, but $x \neq y$.
   (d) $z = \dfrac{x + y}{xy}$, $x$, $y$ arbitrary, but $xy \neq 0$.

3. If $u$ is a differentiable function of $x$, $y$, $z$, prove that
$$d \log u = \frac{du}{u} \qquad \text{and } d \sin u = \cos u \, du.$$

4. If $r = (x^2 + y^2 + z^2)^{1/2}$, show that
$$x\,d\left(\frac{x}{r}\right) + y\,d\left(\frac{y}{r}\right) + z\,d\left(\frac{z}{r}\right) = 0$$
if $r \neq 0$, for all values of $dx$, $dy$, $dz$.

5. Work out $du$ in two ways in each case: once by calculating $\partial u/\partial x$, $\partial u/\partial y$, $\partial u/\partial z$ separately, and once by direct use of formulas for differentials, without conscious use of partial differentiation.
   (a) $u = \dfrac{x^2 z}{y}$.
   (b) $u = \log \dfrac{x}{\sqrt{y^2 + z^2}}$.
   (c) $u = \tan^{-1} \dfrac{xy}{z}$.
   (d) $u = \left(\dfrac{x}{y + z}\right)^{1/2}$.
   (e) $u = \sqrt{x^3 - y^3 - xyz}$.
   (f) $u = e^{xy} \cos xyz$.

6. A wooden box has inside dimensions 2.5 feet by 6 feet by 1.5 feet. If all six faces of the box are 0.5 inch thick, what is the approximate volume of the wood?

7. A quantity $z$ is to be calculated from the formula $z = xy^3 - 8x^2y^2$. Assuming that $x = 1$ with a possible error $\pm 0.01$ and $y = 8$ with a possible error $\pm 0.02$, use differentials to calculate approximately the maximum possible error in $z$.

8. (a) By approximately what per cent might the volume of a right circular cylinder change if the radius of the base were changed by 0.5% and the altitude were changed by 1.5%? (b) What approximate percentage change in the volume would result from a 2% increase in the altitude and a 1.5% decrease in the radius of the base?

9. If $e = \sqrt{a^2 + b^2}/a$, find $de/e$, assuming that $a$ and $b$ are independent variables. If $a = 4$ with a possible error of $\pm 2\%$, and $b = 3$ with a possible error of $\pm 3\%$, what is approximately the greatest possible percentage error in $e$?

10. (a) If $z = 8x^{1/2} \csc \theta$, find $dz$ in terms of $dx$ and $d\theta$ when $x = 2$, $\theta = \pi/4$. (b) What are the maximum possible values of $|dz|$ and $|dz/z|$ if $|dx| \leq 0.1$ and $|d\theta| \leq 0.002$?

11. A fence 4 feet high runs parallel to the wall of a building and 3 feet from it. A man standing at a window in the building looks directly over the fence at a point $P$ on the ground. The man's eyes are 8 feet above the ground. If the fence were 3 inches higher and 4 inches farther from the house, approximately how much farther would the new point $P$ be (a) from the house, (b) from the man's eyes?

12. (a) If $f(x, y) = \sqrt{|xy|}$, what are the values of

$$\left(\frac{\partial f}{\partial x}\right)_{(0,0)} \quad \text{and} \quad \left(\frac{\partial f}{\partial y}\right)_{(0,0)}?$$

(b) Show that the condition (6) is not satisfied at $x = 0$, $y = 0$.

13. If $a \geq 0$ and $b \geq 0$, then $\sqrt{a^2 + b^2} \leq a + b \leq \sqrt{2}\sqrt{a^2 + b^2}$. To prove this, we can show separately that $a^2 + b^2 \leq (a + b)^2$ and $(a + b)^2 \leq 2(a^2 + b^2)$. Explain why each of these latter inequalities is true. Observe that the second one is equivalent to $2ab \leq a^2 + b^2$. Why is this true?

## 19-4 Partial Derivatives of Higher Order

For a function $f(x, y)$ of two variables there are four possible ways in which a second derivative may arise. The notations for these derivatives are as follows:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}, \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\, \partial x},$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\, \partial y}, \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}.$$

It can be proved that if all the first and second derivatives of $f$ are defined at and near $(a, b)$ and if $(\partial^2 f/\partial x\, \partial y)$ and $(\partial^2 f/\partial y\, \partial x)$ are continuous at $(a, b)$, then their values are the same at $(a, b)$. This proof is given in texts on advanced calculus.

**Example 1:** Verify the equality of

$$\frac{\partial^2 f}{\partial x\, \partial y} \quad \text{and} \quad \frac{\partial^2 f}{\partial y\, \partial x}$$

if $f(x, y) = \sqrt{x^2 + y^2} + \frac{y}{x}$, assuming $x \neq 0$.

We have

$$\frac{\partial f}{\partial x} = x(x^2 + y^2)^{-1/2} - yx^{-2}, \qquad \frac{\partial f}{\partial y} = y(x^2 + y^2)^{-1/2} + x^{-1},$$

$$\frac{\partial^2 f}{\partial y\, \partial x} = -xy(x^2 + y^2)^{-3/2} - x^{-2}, \qquad \frac{\partial^2 f}{\partial x\, \partial y} = -xy(x^2 + y^2)^{-3/2} - x^{-2}.$$

The notation for derivatives of third order is almost self-explanatory. Thus

$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y^2}\right) = \frac{\partial^3 f}{\partial x\, \partial y^2}, \qquad \frac{\partial^2}{\partial y^2}\left(\frac{\partial^3 f}{\partial x\, \partial y^2}\right) = \frac{\partial^5 f}{\partial y^2\, \partial x\, \partial y^2},$$

and so on.

It is frequently convenient to have a subscript notation for partial derivatives. If $f$ is a function of $x$ and $y$, we consider $x$ as variable number one and $y$ as variable number two. We write $f_1(x, y)$ for the value of $\partial f/\partial x$ at $(x, y)$, and $f_2(x, y)$ as the value of $\partial f/\partial y$ at $(x, y)$. We then denote $\partial f_1/\partial x$ and $\partial f_1/\partial y$ by $f_{11}(x, y)$ and $f_{12}(x, y)$, respectively, while $\partial f_2/\partial x$ and $\partial f_2/\partial y$ are denoted by $f_{21}(x, y)$ and $f_{22}(x, y)$. Similar notations are employed for derivatives of higher order and for functions of more variables. Thus, in the case of $F(x, y, z)$, $F_{132}(a, b, c)$ would be the value of $(\partial^3 F/\partial y\, \partial z\, \partial x)$ at $x = a$, $y = b$, $z = c$.

The subscript notation for partial derivatives corresponds to the prime notation for ordinary derivatives, where $f'(x)$ denotes the value of the derivative at $x$.

**Example 2:** Let $g$ and $h$ be twice-differentiable functions of a single variable, and let

$$f(x, y) = g(x + 2y) + h(3x^2 - 4y).$$

Obtain the values of $f_1$ and $f_{12}$ in terms of values of $g'$, $g''$, $h'$, and $h''$.

We can use the composite function theorem (chain rule) of § 3-3. Treating $y$ as a constant and differentiating with respect to $x$, we have

$$\frac{\partial}{\partial x} g(x + 2y) = g'(x + 2y) \frac{\partial}{\partial x} (x + 2y) = g'(x + 2y),$$

$$\frac{\partial}{\partial x} h(3x^2 - 4y) = h'(3x^2 - 4y) \frac{\partial}{\partial x} (3x^2 - 4y) = 6xh'(3x^2 - 4y).$$

Hence $$f_1(x, y) = g'(x + 2y) + 6xh'(3x^2 - 4y).$$

Likewise, differentiating this result with respect to $y$, we obtain

$$\begin{aligned} f_{12}(x, y) &= g''(x + 2y) \frac{\partial}{\partial y} (x + 2y) + 6xh''(3x^2 - 4y) \frac{\partial}{\partial y} (3x^2 - 4y) \\ &= 2g''(x + 2y) - 24xh''(3x^2 - 4y). \end{aligned}$$

## EXERCISES

1. If $f(x, y) = \log \sqrt{(x - a)^2 + (y - b)^2}$, show that $\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} = 0$ if $(x, y) \neq (a, b)$.

2. If $F(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$, show that $\dfrac{\partial^2 F}{\partial x^2} + \dfrac{\partial^2 F}{\partial y^2} + \dfrac{\partial^2 F}{\partial z^2} = 0.$

3. If $f(x, t) = e^{x - at} \cos (x - at)$, show that $\dfrac{\partial^2 f}{\partial t^2} = a^2 \dfrac{\partial^2 f}{\partial x^2}$.

4. If $f(x, y) = x^2 \tan^{-1} \dfrac{y}{x} - y^2 \tan^{-1} \dfrac{x}{y}$, show that $\dfrac{\partial^2 f}{\partial y\, \partial x} = \dfrac{x^2 - y^2}{x^2 + y^2}$ if $xy \neq 0$. Our formula defining $f$ is applicable only if $xy \neq 0$. If either $x$ or $y$ is 0 let us define $f(x, y) = 0$. Then it can be shown that $f_{12}(0, 0) = -1$ and $f_{21}(0, 0) = 1$.

5. If $f(x, y) = xy \dfrac{x^2 - y^2}{x^2 + y^2}$, show that

$$\frac{\partial^2 f}{\partial y\, \partial x} = \frac{x^6 + 9x^4y^2 - 9x^2y^4 - y^6}{(x^2 + y^2)^3} \qquad \text{if } x^2 + y^2 \neq 0.$$

6. If $F(x, t) = f(x - at) + g(x + at)$, where $f$ and $g$ are twice-differentiable functions of a real variable, show that $\partial^2 F/\partial t^2 = a^2\, \partial^2 F/\partial x^2$.

7. If $u = f(r)$ and $r = \sqrt{x^2 + y^2}$, show that $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = \dfrac{d^2 u}{dr^2} + \dfrac{1}{r}\dfrac{du}{dr}$. What is the analogous result if $w = F(r)$ and $r = \sqrt{x^2 + y^2 + z^2}$?

8. If $w = F(r)$ and $r = \sqrt{x^2 + y^2 + z^2}$, show that

$$\left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2 = [F'(r)]^2.$$

9. Suppose $z = F(u)$ and $u = f(x, y)$. Express $\dfrac{\partial^2 z}{\partial x^2} + \dfrac{\partial^2 z}{\partial y^2}$ in terms of $F'$, $F''$, and the partial derivatives of $f$.

**10.** In economic theory it is sometimes assumed that the utility to an individual of amounts $x$, $y$, respectively, of two consumers' goods is of the form $u = F(z)$, where $z = f(x, y)$ is a known function of $x$ and $y$, but the function $F$ is unknown in character except that $F'(z) > 0$. The marginal utilities of the two goods are defined as $\partial u/\partial x$, $\partial u/\partial y$. Show that the ratio of these marginal utilities is quite independent of the function $F$.

## 19-5 The Chain Rule

Suppose $u = F(x, y, z)$, where $F$ is differentiable for all considered values of $(x, y, z)$. Let $x, y, z$ be replaced by functions of the independent variables $s$, $t$, say

$$x = f(s, t), \qquad y = g(s, t), \qquad z = h(s, t),$$

and suppose that $f$, $g$, $h$ can each be differentiated partially with respect to $s$ and $t$. Then $u$ becomes a function of $s$ and $t$:

$$u = G(s, t) = F[f(s, t), g(s, t), h(s, t)].$$

Our object is to show that the function $G$ can be differentiated partially with respect to $s$ and $t$. The formulas are

$$\frac{\partial G}{\partial s} = \frac{\partial F}{\partial x}\frac{\partial f}{\partial s} + \frac{\partial F}{\partial y}\frac{\partial g}{\partial s} + \frac{\partial F}{\partial z}\frac{\partial h}{\partial s}, \tag{1}$$

and a similar formula with $t$ in place of $s$. Formula (1) expresses what is called the *chain rule*. The chain rule for functions of one variable was stated in § 3-3, Theorem 3-E.

To prove (1) we must use the condition which expresses the fact that $F$ is differentiable. We fix values of $s$ and $t$. Let the resulting values of the functions $f$, $g$, $h$ be $x$, $y$, $z$. Now suppose $\Delta s \neq 0$ and consider the changes $\Delta x$, $\Delta y$, $\Delta z$, where, for example,

$$\Delta x = f(s + \Delta s, t) - f(s, t).$$

Then let $\qquad \Delta u = F(x + \Delta x, y + \Delta y, z + \Delta z) - F(x, y, z).$

Then also $\qquad \Delta u = G(s + \Delta s, t) - G(s, t),$

and
$$\frac{\partial G}{\partial s} = \lim_{\Delta s \to 0} \frac{\Delta u}{\Delta s}, \qquad \frac{\partial f}{\partial s} = \lim_{\Delta s \to 0} \frac{\Delta x}{\Delta s}, \qquad \text{etc.} \tag{2}$$

Now, if we apply formula (10) of § 19-3 with $\Delta x$, $\Delta y$, $\Delta z$ in place of $dx$, $dy$, $dz$, we have, after dividing by $\Delta s$,

$$\frac{\Delta u}{\Delta s} = \frac{\partial F}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial F}{\partial y}\frac{\Delta y}{\Delta s} + \frac{\partial F}{\partial z}\frac{\Delta z}{\Delta s} + \epsilon\left(\frac{|\Delta x| + |\Delta y| + |\Delta z|}{\Delta s}\right).$$

When we let $\Delta s$ approach 0 we obtain (1) as a result, in view of (2) and the fact that $\epsilon \to 0$. We observe that $\Delta s \to 0$ implies that

$$|\Delta x| + |\Delta y| + |\Delta z| \to 0$$

and $$\left|\frac{|\Delta x| + |\Delta y| + |\Delta z|}{\Delta s}\right| \to \left|\frac{\partial f}{\partial s}\right| + \left|\frac{\partial g}{\partial s}\right| + \left|\frac{\partial h}{\partial s}\right|.$$

A chain rule formula such as (1) is often written entirely in terms of variables instead of with functional symbols. Thus (1) might be written in the form

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial s}.$$

It is necessary to be very clear about the meaning of a letter at each occurrence in a formula such as this. For example, on the left side $u$ is considered a function of $s$ and $t$, while on the right it is considered a function of $x$, $y$, $z$. Likewise, in the symbol $\partial u/\partial x$, $x$ denotes an independent variable, while in the symbol $\partial x/\partial s$, $x$ denotes a dependent variable.

Observe that $x$, $y$, $z$ were the original independent variables in $u = F(x, y, z)$. We shall call them variables of the *first class*. But when we set $x = f(s, t)$, and so on, we introduce new variables $s$, $t$, called variables of the *second class*. In the general form of the chain rule there may be any number of variables in each class. A formula such as (1) contains as many products as there are variables of the first class, and there is a formula like (1) for each variable of the second class.

One of the important consequences of the chain rule, from a theoretical point of view, shows up in the proof of the following statement: *If $u = F(x, y, z)$ is a differentiable function of $x$, $y$, $z$, and if $x$, $y$, $z$ are in turn differentiable functions of other variables, say $s$ and $t$, then $u$ becomes a differentiable function of $s$ and $t$, and the formula*

$$du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy + \frac{\partial u}{\partial z}\,dz$$

*still holds, even when $du$, $dx$, $dy$, $dz$ are all expressed in terms of $ds$ and $dt$ as independent variables.*

In practice the important uses of the chain rule for partial derivatives are those in which at least some of the functions involved are not explicitly given. Quite often the problem of interest is that of calculating the effect on some expression when new variables are introduced.

**Example 1:** If $u = f(x, t)$, calculate the effect upon the expression $\dfrac{\partial^2 u}{\partial t^2} - a^2\dfrac{\partial^2 u}{\partial x^2}$ of letting new variables $p$, $q$ be introduced by setting $p = x - at$, $q = x + at$.

We regard $p$, $q$ as variables of the first class, and $x$, $t$ as variables of the second class. Then

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial p}\frac{\partial p}{\partial x} + \frac{\partial u}{\partial q}\frac{\partial q}{\partial x} = \frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}. \tag{3}$$

Likewise, $$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial p}\frac{\partial p}{\partial t} + \frac{\partial u}{\partial q}\frac{\partial q}{\partial t} = -a\frac{\partial u}{\partial p} + a\frac{\partial u}{\partial q}. \tag{4}$$

At the next step we regard $\partial u/\partial p$ and $\partial u/\partial q$ just as we originally regarded $u$: as functions of $x$ and $t$ through the intermediaries $p$, $q$. Thus we have

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial p}\right) = \frac{\partial}{\partial p}\left(\frac{\partial u}{\partial p}\right)\frac{\partial p}{\partial x} + \frac{\partial}{\partial q}\left(\frac{\partial u}{\partial p}\right)\frac{\partial q}{\partial x} \tag{5}$$

and other similar formulas. This particular formula becomes

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial p}\right) = \frac{\partial^2 u}{\partial p^2} + \frac{\partial^2 u}{\partial q\,\partial p}. \tag{6}$$

Now, from the earlier results (3) and (4) we see that

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial p}\right) + \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial q}\right),$$

$$\frac{\partial^2 u}{\partial t^2} = -a\frac{\partial}{\partial t}\left(\frac{\partial u}{\partial p}\right) + a\frac{\partial}{\partial t}\left(\frac{\partial u}{\partial q}\right).$$

After working out formulas analogous to (6), we ultimately find

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial p^2} - a^2\frac{\partial^2 u}{\partial q\,\partial p} - a^2\frac{\partial^2 u}{\partial p\,\partial q} + a^2\frac{\partial^2 u}{\partial q^2},$$

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial p^2} + \frac{\partial^2 u}{\partial q\,\partial p} + \frac{\partial^2 u}{\partial p\,\partial q} + \frac{\partial^2 u}{\partial q^2}.$$

Consequently, if we assume continuity of the second derivatives, we see that

$$\frac{\partial^2 u}{\partial t^2} - a^2\frac{\partial^2 u}{\partial x^2} = -4a^2\frac{\partial^2 u}{\partial p\,\partial q}.$$

This is the final result of our calculations.

It sometimes happens that a problem involves several variables and several relations between them. For example, if we have a rectangle of length $x$ and width $y$, its area $A$ and perimeter $P$ are given by

$$A = xy, \qquad P = 2x + 2y. \tag{7}$$

Of the four variables $A$, $P$, $x$, $y$, just two are independent, and the other two are then dependent. We could choose $x$ and $P$ as independent. Then $A$ and $y$ would be expressible in the form

$$A = \tfrac{1}{2}xP - x^2, \qquad y = \tfrac{1}{2}P - x. \tag{8}$$

A notation such as $\partial A/\partial x$ is ambiguous, for it does not in itself show whether $A$ is regarded as a function of $x$ and $y$ or as a function of $x$ and $P$. The ambiguity can be removed by a proper use of functional notation. For example, we can write $A = F(x, y)$ when $x$ and $y$ are independent, and $A = G(x, P)$ when $x$ and $P$ are independent. Then $\partial F/\partial x$ and $\partial G/\partial x$ are unambiguous symbols. Another way of removing ambiguity is to write

$$\left(\frac{\partial A}{\partial x}\right)_y \text{ for } \frac{\partial F}{\partial x} \quad \text{and} \quad \left(\frac{\partial A}{\partial x}\right)_P \text{ for } \frac{\partial G}{\partial x}.$$

The presence of $\partial x$ indicates that $x$ is one independent variable, and the literal subscript on the parentheses indicates the other independent variable.

**Example 2:** In the foregoing context show without explicit use of (7) and (8) that

$$\left(\frac{\partial A}{\partial x}\right)_P = \left(\frac{\partial A}{\partial x}\right)_y + \left(\frac{\partial A}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_P. \tag{9}$$

We regard $x$ and $y$ as variables of the first class, $x$ and $P$ as variables of the second class. The connecting equations are $x = x$ and $y = f(x, P)$, the latter standing for the second equation in (8). Then

$$\left(\frac{\partial x}{\partial x}\right)_P = 1 \quad \text{and} \quad \left(\frac{\partial y}{\partial x}\right)_P = \frac{\partial f}{\partial x}. \tag{10}$$

The chain rule is

$$\left(\frac{\partial A}{\partial x}\right)_P = \left(\frac{\partial A}{\partial x}\right)_y\left(\frac{\partial x}{\partial x}\right)_P + \left(\frac{\partial A}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_P,$$

and this is exactly (9), because of the first equation in (10).

## EXERCISES

In all these exercises assume that the functions introduced have continuous derivatives of as many orders as are implied by the context.

**1.** In each case here the variables of the second class are $s$, $t$. Find $\partial u/\partial s$ and $\partial u/\partial t$ (assuming $s$ and $t$ are independent) without first explicitly computing $u$ as a function of $s$ and $t$.

(a) $u = x^2 + xy - y^2$, $x = 2s + t$, $y = s - 3t$.

(b) $u = \dfrac{x - y}{1 + xyz}$, $x = 2s + 3t$, $y = 3s - 4t$, $z = t$.

(c) $u = (x^2 + y^2 + z^2)^{1/2}$, $x = s \cos t$, $y = s \sin t$, $z = st$.

**2.** If $u = F(x, y, z)$ becomes $u = G(r, \theta)$ when $x = r\cos\theta$, $y = r\sin\theta$, $z = r$, express $\partial G/\partial r$, $\partial F/\partial\theta$, and $\partial^2 G/\partial r^2$ in terms of $r$, $\theta$, and the partial derivatives of $F$.

**3.** Suppose $u = F(x, y, z)$, and let $x = r\sin\phi\cos\theta$, $y = r\sin\phi\sin\theta$, $z = r\cos\phi$. If $r = 4$, $\theta = \pi/3$, and $\phi = \pi/6$, the values of $x$, $y$, $z$ are 1, $\sqrt{3}$, $2\sqrt{3}$, respectively. Suppose that

$$\frac{\partial F}{\partial x} = \sqrt{3}, \qquad \frac{\partial F}{\partial y} = 2, \qquad \frac{\partial F}{\partial z} = 1,$$

$$\frac{\partial^2 F}{\partial x^2} = 4, \qquad \frac{\partial^2 F}{\partial y\, \partial x} = \sqrt{3}, \qquad \frac{\partial^2 F}{\partial z\, \partial x} = 2,$$

$$\frac{\partial^2 F}{\partial y^2} = 4, \qquad \frac{\partial^2 F}{\partial z\, \partial y} = 2, \qquad \frac{\partial^2 F}{\partial z^2} = -1,$$

when $x = 1$, $y = \sqrt{3}$, and $z = 2\sqrt{3}$. Find the values of $\partial u/\partial \phi$ and $\partial^2 u/\partial r\, \partial \phi$ (with $r$, $\phi$, $\theta$ independent) when $r = 4$, $\theta = \pi/3$, and $\phi = \pi/6$.

**4.** If $u = F(x, y, z)$, calculate the effect upon the expression $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2}$ of letting new variables $p$, $q$ be introduced by setting

$$p = x + y + z, \qquad q = 2x + y - z.$$

**5.** Suppose $F(x, y)$ becomes $G(r, \theta)$ when we let $x = r\cos\theta$, $y = r\sin\theta$.

(a) Show that $\left(\dfrac{\partial G}{\partial r}\right)^2 + \dfrac{1}{r^2}\left(\dfrac{\partial G}{\partial \theta}\right)^2 = \left(\dfrac{\partial F}{\partial x}\right)^2 + \left(\dfrac{\partial F}{\partial y}\right)^2.$

(b) Show that $\dfrac{\partial^2 G}{\partial r^2} + \dfrac{1}{r}\dfrac{\partial G}{\partial r} + \dfrac{1}{r^2}\dfrac{\partial^2 G}{\partial \theta^2} = \dfrac{\partial^2 F}{\partial x^2} + \dfrac{\partial^2 F}{\partial y^2}.$

**6.** If $w = f(x^2 - y^2, y^2 - x^2)$, show that

$$y\frac{\partial w}{\partial x} + x\frac{\partial w}{\partial y} = 0.$$

*Suggestion:* Write $w = f(u, v)$, $u = x^2 - y^2$, $v = y^2 - x^2$.

**7.** If $w = F\left(\dfrac{y - x}{xy}, \dfrac{z - y}{yz}\right)$, show that

$$x^2\frac{\partial w}{\partial x} + y^2\frac{\partial w}{\partial y} + z^2\frac{\partial w}{\partial z} = 0.$$

**8.** Suppose that variables $x$, $y$, $z$, $u$, $v$ are related as follows:

$$u = F(x, y, z), \qquad z = f(x, y, v).$$

Let $F[x, y, f(x, y, v)] = G(x, y, v)$. Then we might write

$$\left(\frac{\partial u}{\partial x}\right)_{yz} = \frac{\partial F}{\partial x}, \qquad \left(\frac{\partial u}{\partial x}\right)_{yv} = \frac{\partial G}{\partial x},$$

with other similar notations to avoid ambiguity, much as in Example 2. Show that

$$\frac{\partial G}{\partial x} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial f}{\partial x}$$

and
$$\left(\frac{\partial u}{\partial y}\right)_{xv} = \left(\frac{\partial u}{\partial y}\right)_{xz} + \left(\frac{\partial u}{\partial z}\right)_{xy}\left(\frac{\partial z}{\partial y}\right)_{xv}.$$

What is the corresponding formula for $\partial G/\partial v$? Verify these results in the particular case

$$u = x^2 + y^2 + z^2, \qquad v = xyz.$$

**9.** (a) Deduce formula (9) by use of differentials, as follows:

$$dA = \left(\frac{\partial A}{\partial x}\right)_y dx + \left(\frac{\partial A}{\partial y}\right)_x dy, \qquad dA = \left(\frac{\partial A}{\partial x}\right)_P dx + \left(\frac{\partial A}{\partial P}\right)_x dP,$$

$$dy = \left(\frac{\partial y}{\partial x}\right)_P dx + \left(\frac{\partial y}{\partial P}\right)_x dP.$$

Now substitute for $dy$ from the third into the first formula and compare the result with the second formula. Equating coefficients of $dx$ gives (9). What is obtained by equating coefficients of $dP$?

(b) Use the method of (a) to obtain the results sought in Exercise 8.

**10.** Let

$$D = \begin{vmatrix} a & b & c \\ x & y & z \\ u & v & w \end{vmatrix}$$

where $a, b, c, x, y, z, u, v, w$ are functions of $t$. Using the chain rule and the rules for expansion of the determinant by minors of rows, show that

$$\frac{dD}{dt} = \begin{vmatrix} a' & b' & c' \\ x & y & z \\ u & v & w \end{vmatrix} + \begin{vmatrix} a & b & c \\ x' & y' & z' \\ u & v & w \end{vmatrix} + \begin{vmatrix} a & b & c \\ x & y & z \\ u' & v' & w' \end{vmatrix},$$

where primes denote differentiation with respect to $t$.

## 19-6 Extreme Value Problems

Consider a function $f$ of the two variables $x, y$. A point $(a, b)$ in the domain of definition of $f$ is called an *interior point* if there is some circle with center at $(a, b)$ such that all points inside the circle are also in the domain of definition of $f$. A point of the domain which is not an interior point is said to be on the *boundary* of the domain.

**Example 1:** Let $f(x, y) = \sqrt{y - x^2}$. Then the domain of definition of $f$ is made up of the points $(x, y)$ for which $y \geq x^2$. Those for which $y > x^2$ are interior points. Those for which $y = x^2$ are on the boundary.

Let $D$ be the domain of definition of $f$. If there is some point $(a, b)$ in $D$ such that $f(x, y) \leq f(a, b)$ for every point $(x, y)$ in $D$, we say that $f$ attains an *absolute maximum* at $(a, b)$. In contrast to the notion of an absolute maximum we have the notion of a *relative maximum*, which is defined as follows: $f$ attains a relative maximum at $(a, b)$ if there is some circle with center at $(a, b)$ such that $f(x, y) \leq f(a, b)$ for every point $(x, y)$ of $D$ which is inside this circle.

The notions of absolute and relative minimum values are defined in a similar way, with $f(x, y) \geq f(a, b)$ instead of $f(x, y) \leq f(a, b)$.

By an *extreme value* we mean either a maximum or a minimum value.

The following theorem is important in the study of extreme values of functions.

THEOREM 19-B. *Suppose that f attains a relative extreme value at the point (a, b) in the domain D where f is defined. Suppose that (a, b) is an interior point of D and that f has first partial derivatives at (a, b). Then*

$$\frac{\partial f}{\partial x} = 0 \quad \textit{and} \quad \frac{\partial f}{\partial y} = 0 \quad \textit{at } (a, b). \tag{1}$$

*Proof.* This theorem is for functions of two variables what Theorem 2-B (in § 2-1) is for functions of one variable. We can base the proof on Theorem 2-B. Consider $f(x, b)$, in which $x$ alone is variable. This function of $x$ is defined for $x$ in an interval which contains $x = a$ and extends on either side of $x = a$ [because $(a, b)$ is an interior point of $D$]. The function is differentiable with respect to $x$ at $x = a$, and the function has a relative extreme there. Hence, by Theorem 2-B,

$$\frac{d}{dx} f(x, b) = 0 \quad \text{at} \quad x = a.$$

This is the same as saying that $\partial f/\partial x = 0$ at $(a, b)$. The assertion about $\partial f/\partial y$ is proved in the same way.

Two comments should be made at once. The assumption that $(a, b)$ is an interior point of $D$ is essential. And the conditions

$$\frac{\partial f}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} = 0 \tag{2}$$

may be satisfied at points where $f$ does not have a relative extreme. If we think of cases in which $f$ is differentiable, and visualize the surface $z = f(x, y)$ as a graphical representation of the function, the legitimacy of the remarks is quickly evident. The conditions (2) at a point mean that the tangent plane at that point is horizontal (perpendicular to the $z$-axis). Now a horizontal tangent plane does not always indicate a relative extreme value of $z$. Consider the hyperbolic paraboloid $z = y^2 - x^2$ at $(0, 0)$ for instance (see Fig. 18-15). Also, if there is a relative extreme at a point on the boundary of $D$, rather than at an interior point, then the tangent plane at the point need not be horizontal. For instance, suppose that $f(x, y)$ is defined to be $1 - x - y$ when $x \geq 0$ and $y \geq 0$, and is not defined unless $x$ and $y$ satisfy these conditions. Then $f$ attains the absolute maximum value 1 at $(0, 0)$, which is a boundary point of the domain. In this case conditions (2) are never satisfied.

A point at which the first partial derivatives of a function all exist and are equal to zero is called a *critical point* of the function. If, at some point of the domain of $f$, at least one of the partial derivatives fails to be defined, then we call this point a *singular point* of the function. (This definition of a singular point is made purely for convenience in our present discussion, and is not meant to apply beyond this discussion.)

Now suppose we have a certain function, and suppose we know that it does attain an absolute maximum value. How shall we go about it to find the point or points at which the absolute maximum is attained? From Theorem 19-B we can infer that the maximum occurs either (a) at an interior critical point, or (b) at an interior singular point, or (c) at a point on the boundary. In many problems the alternatives (b) and (c) can be ruled out for one reason or another, and we are left with (a). The procedure is then to locate the interior critical points. If there is only one such point, it must be the point we seek. If there are several interior critical points, we must compute the function values at each of them and decide which one is the sought maximum.

The analysis is of exactly the same kind for minimum problems.

**Example 2:** Find by calculus the point of the plane $3x + 4y - z = 26$ which is nearest the origin.

The square of the distance from the origin to $P(x, y, z)$ is $u = x^2 + y^2 + z^2$. We are to make $u$ a minimum when $x$, $y$, $z$ are related by the equation of the plane. Hence we seek the absolute minimum of the function

$$u = f(x, y) = x^2 + y^2 + (3x + 4y - 26)^2.$$

The domain of $f$ is the entire $xy$-plane. In this case all points are interior points and there are no singular points. Since a minimum certainly does exist (we take this for granted), we locate it by searching for critical points. We write

$$\frac{\partial f}{\partial x} = 2x + 6(3x + 4y - 26) = 0,$$

$$\frac{\partial f}{\partial y} = 2y + 8(3x + 4y - 26) = 0.$$

Simplifying, we obtain the simultaneous linear equations

$$5x + 6y = 39, \qquad 12x + 17y = 104.$$

The solution is $x = 3$, $y = 4$. The corresponding value of $z$ is $9 + 16 - 26 = -1$. Since there is just the one critical point, it must furnish the sought minimum. The required point on the plane is $(3, 4, -1)$.

The theory of finding extreme values of functions of more than two variables is the same in general outline as for two variables. The problem of solving simultaneous equations to find critical points may lead to great complications. In this book we do not attempt to cope with all the possible intricacies of extreme value problems. More details, both of theory and technique, may be found in books on advanced calculus.

*A Second Derivative Test for Extreme Values*

When several interior critical points are found, it is sometimes possible to use tests by second derivatives to distinguish between points of relative maximum, points of relative minimum, and critical points which furnish neither a maximum nor a minimum. Such tests are not always convenient or necessary in extremal problems, however.

Without giving a complete justification of these tests by second derivatives, we can nevertheless make them plausible by the following considerations. The tests will be stated presently. Consider the function

$$z = f(x, y) = Ax^2 + 2Bxy + Cy^2. \tag{3}$$

It has a critical point at (0, 0), as we can easily verify. Let us assume that $B^2 - AC \neq 0$. Then the following assertions are true: (a) if $B^2 - AC < 0$ and $A > 0$, $f$ has a relative minimum at (0, 0); (b) if $B^2 - AC < 0$ and $A < 0$, $f$ has a relative maximum at (0, 0); (c) if $B^2 - AC > 0$, $f$ has neither a maximum nor a minimum at (0, 0).

To justify these assertions, we use the results on homogeneous quadratic forms from §§ 7-7, 7-8. By a rotation of coordinate axes in the $xy$-plane we can bring (3) to the form

$$z = ax'^2 + cy'^2 \tag{4}$$

where the new coordinates are $x'$, $y'$. Moreover,

$$B^2 - AC = -ac, \qquad A + C = a + c. \tag{5}$$

Now, if $B^2 - AC < 0$, we have $ac > 0$, so that $a$ and $c$ are of the same sign. Moreover, $A$ and $C$ are of the same sign, for $AC < 0$ would imply $B^2 - AC > 0$. Hence in this case (4) is the equation of a paraboloid with vertex at (0, 0, 0). It opens upward if $a$ and $c$ are positive, i.e., if $A > 0$, and this makes $z$ a minimum at $x = y = 0$. It opens downward, making $z$ a maximum, if $a$ and $c$ are negative, i.e., if $A < 0$. When $B^2 - AC > 0$, $a$ and $c$ are of opposite signs, and (4) represents a hyperbolic paraboloid, with $z$ having neither a maximum nor a minimum at $x = y = 0$.

The second derivative tests we referred to are stated as follows:

THEOREM 19-C. *Suppose that $f$ is defined and differentiable throughout a domain of definition of which $(a, b)$ is an interior point. Let $(a, b)$ be a critical point of $f$ and suppose that $\partial f/\partial x$ and $\partial f/\partial y$ are differentiable at $(a, b)$. Let*

$$A = f_{11}(a, b), \qquad B = f_{12}(a, b), \qquad C = f_{22}(a, b). \tag{6}$$

*Then, if $B^2 - AC < 0$,*

(a) *$f$ has a relative minimum at $(a, b)$ if $A > 0$;*

(b) *$f$ has a relative maximum at $(a, b)$ if $A < 0$.*

*If $B^2 - AC > 0$, $f$ has neither a relative maximum nor a relative minimum* *at $(a, b)$. If $B^2 - AC = 0$, no conclusion at all about relative extreme values can be drawn unless some further information is given.*

The proof of Theorem 19-C can be made by an argument which falls back on what was said in connection with the quadratic form in (3). If we write $x - a = h$, $y - b = k$, then $h$ and $k$ are small when $(x, y)$ is near $(a, b)$. The hypotheses about differentiability make it possible to show that when $h$ and $k$ are sufficiently small, $f(a + h, b + k) - f(a, b)$ has the same sign as

$$Ah^2 + 2Bhk + Ck^2,$$

where $A$, $B$, $C$ are given by (6). All statements except the last one in the theorem follow from this. That nothing can be concluded without further information if $B^2 - AC = 0$ can be shown by citing various examples.

**Example 3:** Investigate the critical points of the function $f(x, y) = xy(12 - 4x - 3y)$. We find

$$\frac{\partial f}{\partial x} = 12y - 8xy - 3y^2 = y(12 - 8x - 3y),$$

$$\frac{\partial f}{\partial y} = 12x - 4x^2 - 6xy = x(12 - 4x - 6y).$$

The critical points (four of them) are given by the solutions of the following pairs of equations:

$$y = 0,\ 12 - 4x - 6y = 0, \qquad \text{solution } (3, 0);$$

$$x = 0,\ 12 - 8x - 3y = 0, \qquad \text{solution } (0, 4);$$

$$y = 0,\ x = 0, \qquad \text{solution } (0, 0);$$

$$\left.\begin{matrix} 12 - 8x - 3y = 0 \\ 12 - 4x - 6y = 0 \end{matrix}\right\}, \qquad \text{solution } (1, \tfrac{4}{3}).$$

Now $$\frac{\partial^2 f}{\partial x^2} = -8y, \quad \frac{\partial^2 f}{\partial x\, \partial y} = 12 - 8x - 6y, \quad \frac{\partial^2 f}{\partial y^2} = -6x.$$

Using Theorem 19-C, we find that $B^2 - AC > 0$ at each of the points $(3, 0)$, $(0, 4)$, $(0, 0)$. Hence none of these points yields a relative extreme. At $(1, \frac{4}{3})$ we find $A = -\frac{32}{3}$, $B = -4$, $C = -6$, so $B^2 - AC = -48$. Since $A < 0$, the function has a relative maximum at $(1, \frac{4}{3})$.

*Existence of Absolute Extreme Values*

In the study of extremal values of functions of several variables it is important to know the appropriate analogue of Theorem 2-A (see § 2-1). In that theorem the hypothesis was that the function of one variable $x$ was continuous at each point of a closed interval on the $x$-axis, *including the end-points* of the interval. If one or both end-points were omitted in the specification, the theorem would no longer be true. Likewise, the finiteness of the interval is also essential. For instance, a function $f$ which is continuous for each $x$ such that $x \geq 0$ need not attain either an absolute minimum or an absolute maximum. One can easily think of examples to justify what has just been said.

Now consider the situation for functions of two variables. We shall explain what is meant by a bounded and closed set of points in the $xy$-plane. A set of points (i.e., a collection of points) is said to be *bounded* if it is entirely contained within some square in the plane. The set of all points $(x, y)$ such that $x^2 + (y - 2)^2 \leq 16$ is bounded. The set of points $(x, y)$ such that $-1 \leq x \leq 4$, with no restriction on $y$, is *not* bounded. A set $S$ of points is said to be *closed* if $S$ contains all points $Q$ such that points of $S$ can be found as close as one pleases to $Q$. The set $S$ of points $(x, y)$ such that $0 < x \leq 1$ and $0 \leq y \leq 1$ is not closed. This is because the points $(0, y)$ with $0 \leq y \leq 1$ have been excluded from the set $S$. For each such point one can find points of $S$ as close as one pleases to it. If these points were included in $S$, the set would be closed. Another example: The set of all points $(x, y)$ such that $x^2 + y^2 \leq 1$ is closed, but it ceases to be closed if we omit from it any finite number of its points.

THEOREM 19-D. *Suppose that $S$ is a bounded and closed set of points in the domain of definition of $f$ (a function of two variables). Let $f$ be continuous at each point of $S$. Then, if we consider the values which $f$ assumes at the points of $S$, there is an absolute maximum of these values, and also an absolute minimum.*

We omit the proof, which belongs to a course in advanced calculus. The example $f(x, y) = xy$ with $S$ the set of points such that $x \geq 0$ shows that it is essential for $S$ to be bounded, for in this case $S$ is closed, though not bounded, and yet the values of $f$ have neither a maximum nor a minimum. The example $f(x, y) = 1/(x^2 + y^2)$, with $S$ composed of all points such that $0 < x^2 + y^2 \leq 1$, shows that it is essential to have $S$ closed. In this example $S$ is bounded but not closed [because $(0, 0)$ is not in $S$], and the values of $f$ have no absolute maximum.

In the next example we show how the use of Theorem 19-D may enable us to avoid the use of second-derivative tests.

**Example 4:** Find the absolute maximum of $f(x, y) = xy(12 - 4x - 3y)$ on the set $S$ consisting of all points $(x, y)$ on or inside the triangle with vertices at $(0, 0)$, $(3, 0)$, $(0, 4)$.

This set is closed and bounded, and $f$ is continuous at each point of $S$. Moreover, we see from the definition of $f$ that $f(x, y) = 0$ along each line forming a side of the triangle, while $f(x, y) > 0$ at points inside the triangle. By Theorem 19-D there does exist an absolute maximum of the values of $f$ on $S$, and it is evident that the maximum does not occur anywhere on the periphery of the triangle. Hence the absolute maximum must occur at an interior point of $S$. Since there are no interior singular points, the absolute maximum must occur at an interior critical point. But, as we can see from the solution of Example 2, the only interior critical point is $(1, \frac{4}{3})$. Hence this must be the point of the absolute maximum value. We reach this conclusion with the aid of Theorem 19-D, without any use of second derivatives.

## EXERCISES

**1.** Find the shortest distance from the point $(0, 2, 1)$ to the plane $4x + y + 4z = 39$.

**2.** Find the maximum possible volume for a rectangular box without a top if the combined area of the four sides and bottom is 108 square feet. Take for granted that the maximum exists.

**3.** A rectangular boxlike enclosure is to be built with Fiberglas paneling covering the top, two ends, and the back. The volume enclosed is to be 3456 cubic feet. What is the least possible square footage of paneling needed? Take for granted that there is an absolute minimum area as required.

**4.** Find the absolute maximum of $f(x, y) = xy(ab - bx - ay)$ in the closed triangular region with vertices at $(0, 0)$, $(a, 0)$, $(0, b)$.

**5.** Find the absolute maximum of $f(x, y) = x^2y(ab - bx - ay)$ on the set $S$ consisting of points $(x, y)$ such that $x \geq 0$, $y \geq 0$.

**6.** Find the absolute maximum value of $f(x, y) = \sin x \sin y \sin (x + y)$ in the closed triangular region with vertices at $(0, 0)$, $(\pi, 0)$, $(0, \pi)$.

**7.** (a) If $x$, $y$, $z$ are positive and such that $xyz = 9$, what is the smallest possible value of $2x + 3y + 4z$? Assume that the absolute minimum exists. (b) If $C$ is the minimum in (a), show that the plane $2x + 3y + 4z = C$ is tangent to the surface $xyz = 9$ at the point $(x, y, z)$ which produces the minimum.

**8.** (a) A manufacturer produces safety razors and blades at a cost of 50 cents per razor and 15 cents per dozen blades. If he charges $x$ cents per razor and $y$ cents per dozen blades, he finds that he can sell $(1944)10^4/x^2y$ razors and $(7776)10^4/xy^2$ dozen blades daily. How should he fix prices so as to maximize his profit? (b) If in (a) we replace $(1944)10^4$ by $A$ and $(7776)10^4$ by $B$, show that the conditions for maximum total profit will result in a profit on razors alone if and only if $A \geq 2B$, and in a profit on blades alone if and only if $2A \leq B$.

**9.** (a) Find positive numbers $x$, $y$, $z$ such that $x + y + z = 24$ and $xyz^2$ is as large as possible. (b) Find positive numbers $x$, $y$, $z$ such that $x + y + z = 18$ and $xy^2z^3$ is as large as possible.

**10.** Find all critical points of each function and state what you can about relative extremes on the basis of Theorem 19-C. A critical point at which

$B^2 - AC > 0$ in Theorem 19-C is called a *saddle point.* A critical point at which $B^2 - AC = 0$ is called *degenerate.*

(a) $f(x, y) = 2x^3 + 16y^3 - 9xy$. (b) $f(x, y) = \frac{xy}{27} + \frac{1}{x} - \frac{1}{y}$.

**11.** Proceed as directed in Exercise 10.
(a) $f(x, y) = 3x^2 - 2xy - 2x + 3y + 1$.
(b) $f(x, y) = 3xy - x^2 - 3y^2 + 7x - 12y - 10$.
(c) $f(x, y) = x^2 + 3y^4 - 4y^3 - 12y^2$.
(d) $f(x, y) = x^4 + y^4 + 32x - 4y + 52$.
(e) $f(x, y) = (3 - x)(2 - y)(4x + 3y - 12)$.
(f) $f(x, y) = x^2y^2 - 5x^2 - 8xy - 5y^2$.

**12.** (a) By analyzing the geometrical meaning of each factor being positive or negative, describe the part of the plane where the function

$$f(x, y) = [x^2 + (y - 2)^2 - 4][x^2 + (y - 1)^2 - 1]$$

is negative, and the part where it is positive.
(b) Does $f$ have a relative extreme at $(0, 0)$?
(c) Consider an arbitrary straight line through $(0, 0)$, and consider the values of $f$ at points on this line. Show that for these values there is a relative minimum at $(0, 0)$.

## 19-7 Directional Derivatives. Gradients

Consider a differentiable function of two variables. We are familiar with the interpretation of $\partial f/\partial x$ as the rate of change of $f(x, y)$ with respect to $x$, or, what is the same thing, the rate of change of $f(x, y)$ per unit distance along the line through $(x, y)$ in the positive $x$-direction. Now we consider a generalization, in which we seek the rate of change of $f(x, y)$ per unit distance in any specified direction from $(x, y)$. We select a direction and consider an arc of any smooth curve starting at $(x, y)$ and going off in this direction. See Fig. 19-4, in which the arc is $PQ$, of length $\Delta s$, and the specified direction from $P$ is indicated by a vector. Let $\alpha$ be the counterclockwise angle from the positive $x$-direction to the specified direction. The point $P$ is $(x, y)$, and $Q$ is $(x + \Delta x, y + \Delta y)$. The rate of change of $f$ along the arc in question at $P$ is the limit

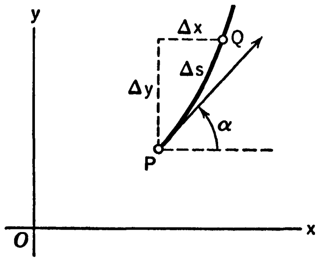



**Fig. 19-4**

$$\lim_{\Delta s \to 0} \frac{f(x + \Delta x, y + \Delta y) - f(x, y)}{\Delta s}. \tag{1}$$

This limit is taken as $Q$ approaches $P$ along the arc. Since $f$ is differentiable, we can write

$$f(x + \Delta x, y + \Delta y) - f(x, y) = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \epsilon\,(|\Delta x| + |\Delta y|),$$

where $\epsilon \to 0$ as $\Delta x$ and $\Delta y$ approach 0. Hence we see that the limit in (1) is the same as the limit

$$\lim_{\Delta s \to 0} \frac{\partial f}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta s} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds}. \tag{2}$$

But
$$\frac{dx}{ds} = \cos\alpha, \qquad \frac{dy}{ds} = \sin\alpha,$$

and so we can write (2) in the form

$$\frac{df}{ds} = \frac{\partial f}{\partial x}\cos\alpha + \frac{\partial f}{\partial y}\sin\alpha. \tag{3}$$

*We call $df/ds$ the directional derivative of $f$ at $(x, y)$ in the direction determined by the angle $\alpha$.* The value of $df/ds$ does not depend on the particular curve, so long as the tangent to the curve at $(x, y)$, drawn in the proper sense, has the direction determined by $\alpha$. Once $\partial f/\partial x$ and $\partial f/\partial y$ have been computed, the value of $df/ds$ is entirely determined by $\alpha$.

**Example 1:** Find the directional derivative of $f(x, y) = x^2y^3/48$ in the direction of inclination $\alpha$ at (2, 3). Discuss the way in which this directional derivative varies as $\alpha$ varies, and show how the results are related to the level curves of the function.

We find

$$\frac{\partial f}{\partial x} = \frac{xy^3}{24} = \frac{9}{4} \quad \text{at } (2, 3),$$

$$\frac{\partial f}{\partial y} = \frac{x^2y^2}{16} = \frac{9}{4} \quad \text{at } (2, 3).$$

Then, for the direction of inclination $\alpha$,

$$\frac{df}{ds} = \frac{9}{4}(\cos\alpha + \sin\alpha).$$

If we plot $df/ds$ as a function of $\alpha$ when $0 \le \alpha \le 2\pi$, the graph is that shown in Fig. 19-5. The easiest way to obtain this graph is to write

$$\cos\alpha + \sin\alpha = \sqrt{2}\left(\cos\alpha \cdot \frac{1}{\sqrt{2}} + \sin\alpha \cdot \frac{1}{\sqrt{2}}\right) = \sqrt{2}\cos\left(\alpha - \frac{\pi}{4}\right).$$

The maximum value of $df/ds$ is $9\sqrt{2}/4$, at $\alpha = \pi/4$; the minimum is $-9\sqrt{2}/4$, at $\alpha = 5\pi/4$. These extremes could also be found by computing that

$$\frac{d}{d\alpha}\left(\frac{df}{ds}\right) = \frac{9}{4}(-\sin\alpha + \cos\alpha) = 0$$

when $\tan\alpha = 1$, i.e., when $\alpha = \pi/4$ or $\alpha = 5\pi/4$.

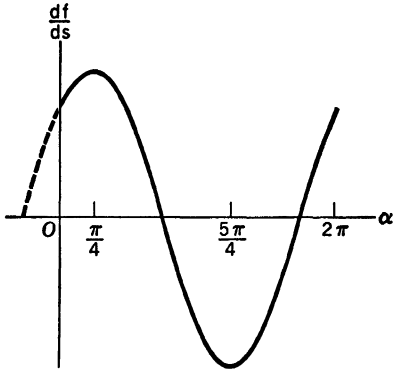



**Fig. 19-5**

The level curves of the function $f$ are the curves $x^2y^3 = 48C$. The level curve through (2, 3) is the one for which $C = \frac{9}{4}$. The slope of this level curve at (2, 3) is given by

$$3x^2y^2\frac{dy}{dx} + 2xy^3 = 0, \qquad \frac{dy}{dx} = -\frac{2xy^3}{3x^2y^2} = -\frac{2y}{3x} = -1.$$

See Fig. 19-6. We observe that at (2, 3), $\alpha = \pi/4$ and $\alpha = 5\pi/4$ give directions perpendicular to the level curve through (2, 3), whereas $\alpha = 3\pi/4$ and $\alpha = 7\pi/4$ give directions tangent to this level curve. Note that at (2, 3),

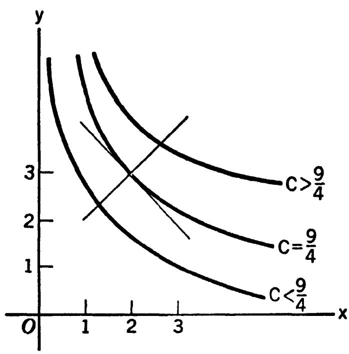



**Fig. 19-6**

$df/ds$ is a maximum for $\alpha = \pi/4$; this is the direction in which $f$ increases most rapidly.

The situation described in the foregoing example illustrates a general principle. Suppose $f$ is a differentiable function and that $(x, y)$ is not a critical point of the function, so that $\partial f/\partial x$ and $\partial f/\partial y$ are not both zero at $(x, y)$. Then, if we draw the level curve through the point $(x, y)$, we shall find that the direction of most rapid increase of $f$ at $(x, y)$ is perpendicular to this level curve, and toward the side on which the values of $f$ are larger than at $(x, y)$. The opposite direction gives the most rapid rate of *decrease* of the function.

*The Gradient of a Function*

We shall now examine the directional derivative concept from a vector standpoint. We are dealing with the $xy$-plane, so we use the notations which were developed in §13-2. In particular, $\mathbf{i}$ and $\mathbf{j}$ are the unit vectors in the $x$-direction and $y$-direction, respectively.

Suppose $f$ is a differentiable function; for a fixed point $(x, y)$ we define a vector called the gradient of $f$ at $(x, y)$, denoted by grad $f$ and defined by

$$\text{grad } f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}. \tag{4}$$

Thus the $x$ and $y$ components of grad $f$ are $\partial f/\partial x$, $\partial f/\partial y$. Also, for an arbitrary $\alpha$ consider the vector

$$\mathbf{T} = \cos\alpha\mathbf{i} + \sin\alpha\mathbf{j}. \tag{5}$$

This vector $\mathbf{T}$ has unit length, and the counterclockwise angle from $\mathbf{i}$ to $\mathbf{T}$ is $\alpha$. Now consider the component of grad $f$ in the direction of $\mathbf{T}$. This component is simply the scalar product

$$\left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}\right)\cdot(\cos\alpha\mathbf{i} + \sin\alpha\mathbf{j}),$$

which on computation turns out to be

$$\frac{\partial f}{\partial x}\cos\alpha + \frac{\partial f}{\partial y}\sin\alpha.$$

This is exactly the directional derivative of $f$ in the direction of $\mathbf{T}$. The gradient of $f$ at $(x, y)$, therefore, is a vector with the property that its component in any specified direction at $(x, y)$ is the directional derivative of $f$ in that direction.

*Gradients in Three Dimensions*

If $F$ is a differentiable function of three variables $x$, $y$, $z$, its gradient at $(x, y, z)$ is defined as the vector

$$\text{grad } F = \frac{\partial F}{\partial x}\mathbf{i} + \frac{\partial F}{\partial y}\mathbf{j} + \frac{\partial F}{\partial z}\mathbf{k}. \tag{6}$$

If $\mathbf{T}$ is a unit vector in some given direction, the component of grad $F$ in the direction of $\mathbf{T}$ is $(\text{grad } F)\cdot\mathbf{T}$. Now, if $\mathbf{T}$ makes angles $\alpha$, $\beta$, $\gamma$ with the coordinate axes of $x$, $y$, $z$, respectively, then

$$\mathbf{T} = \cos\alpha\mathbf{i} + \cos\beta\mathbf{j} + \cos\gamma\mathbf{k}, \tag{7}$$

and so

$$(\text{grad } F)\cdot\mathbf{T} = \frac{\partial F}{\partial x}\cos\alpha + \frac{\partial F}{\partial y}\cos\beta + \frac{\partial F}{\partial z}\cos\gamma. \tag{8}$$

This is the same as the directional derivative of $F$ in the direction of $\mathbf{T}$, because if we have a smooth curve passing through $(x, y, z)$, with arc length $s$ increasing in the direction of $\mathbf{T}$, we know from §18-6 that

$$\cos\alpha = \frac{dx}{ds}, \qquad \cos\beta = \frac{dy}{ds}, \qquad \cos\gamma = \frac{dz}{ds}.$$

**Example 2:** Let a particle of mass $m$ be fixed at a point $(a, b, c)$. According to the law of gravitation, this mass attracts any other mass particle with a force toward $(a, b, c)$ and of magnitude inversely proportional to the square of the distance between the masses. If suitable units are adopted, the function

$$F(x, y, z) = \frac{m}{[(x-a)^2 + (y-b)^2 + (z-c)^2]^{1/2}}$$

is called the *Newtonian potential* of the mass $m$ at $(a, b, c)$. Notice that if $r$ is the distance between $(x, y, z)$ and $(a, b, c)$, then

$$F(x, y, z) = \frac{m}{r}. \tag{9}$$

This is because

$$r^2 = (x-a)^2 + (y-b)^2 + (z-c)^2. \tag{10}$$

We shall show that grad $F$ is a vector of length $m/r^2$ which has the direction of the line from $(x, y, z)$ to $(a, b, c)$. Hence, with suitable units grad $F$ represents the gravitational force which $m$ would exert on a unit mass particle at $(x, y, z)$. The importance of the function $F$ stems from this fact.

To compute grad $F$ it is easiest to use (9) and (10):

$$\frac{\partial F}{\partial x} = -\frac{m}{r^2}\frac{\partial r}{\partial x}, \quad 2r\frac{\partial r}{\partial x} = 2(x-a), \quad \frac{\partial r}{\partial x} = \frac{x-a}{r},$$

whence

$$\frac{\partial F}{\partial x} = -\frac{m}{r^3}(x-a) = \frac{m}{r^2}\frac{(a-x)}{r}.$$

By symmetry, then

$$\text{grad } F = \frac{m}{r^2}\left(\frac{a-x}{r}\mathbf{i} + \frac{b-y}{r}\mathbf{j} + \frac{c-z}{r}\mathbf{k}\right). \tag{11}$$

This formula for grad $F$ admits exactly the interpretation stated previously, for the length of grad $F$ is $m/r^2$, and

$$\frac{a-x}{r}\mathbf{i} + \frac{b-y}{r}\mathbf{j} + \frac{c-z}{r}\mathbf{k}$$

is clearly a vector of unit length in the direction from $(x, y, z)$ toward $(a, b, c)$.

We note, incidentally, that the level surfaces of $F$ are spheres with center at $(a, b, c)$. If we think of grad $F$ as a vector based at $(x, y, z)$, we see that it is perpendicular to the level surface at $(x, y, z)$ and that it points in the direction of increasing values of $F$. See Fig. 19-7.

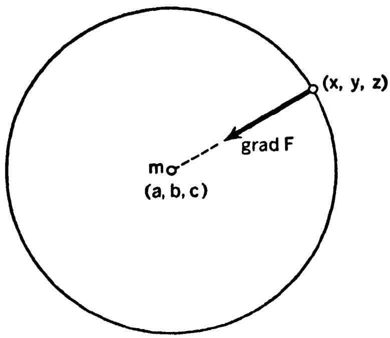



**Fig. 19-7**

Since the directional derivative of a function in any given direction is the component of the gradient in that direction, it is clear that at any particular point the directional derivative is greatest in the direction of the gradient at that point. In order to see that grad $F$ at $(x, y, z)$ is perpendicular to the level surface of $F$ through the point $(x, y, z)$, we need some formulas which are developed in the section following this one. See especially (7) in Theorem 19-E and Exercises 3 and 4 in §19-8.

## EXERCISES

**1.** In each case find the rate of change of the given function at the first indicated point, in the direction toward the second indicated point.

(a) $\log [(x-1)^2 + y^2]^{1/2}$ at $(0, 0)$ toward $(3, 4)$.

(b) $(x-2)^2 + 4(y-1)^2$ at $(4, 2)$ toward $(0, 0)$.

(c) $4x^2 + xy + 9y^2$ at $(1, 2)$ toward $(5, -1)$.

(d) $(x^2 + y^2 + z^2)^{3/2}$ at $(3, 4, -12)$ toward $(15, 7, -8)$.

(e) $xy + yz + zx + xyz$ at $(2, -1, 3)$ toward $(-6, -2, -1)$.

(f) $x(x^2 + y^2 + z^2)^{-1/2}$ at $(8, -1, 4)$ toward $(12, 11, 1)$.

**2.** Find grad $F$ and the maximum possible value of the directional derivative of $F$ at the point indicated.

(a) $F(x, y, z) = x^2 - 3xy + 2y^2 - 4yz + 6z^2$ at $(1, 0, 2)$.

(b) $F(x, y, z) = \dfrac{x^2}{16} + \dfrac{y^2}{25} - \dfrac{z^2}{9}$ at $(4, -5, 3)$.

(c) $F(x, y, z) = e^{-xyz}(x^2 + y^2 + z^2)$ at $(1, 0, 0)$.

(d) $F(x, y, z) = \dfrac{x}{x^2 + y^2} + \dfrac{y}{x^2 + z^2}$ at $(3, 12, 4)$.

3. If $f(x, y) = \dfrac{y}{x^2 + y^2}$, find the rate of change of $f$: (a) at the point (1 ,2) in a direction making an angle of 120° with the positive $x$-axis; (b) at the point (0, 3) in the direction of the vector $-12\mathbf{i} - 5\mathbf{j}$; (c) at the point $(x, y)$ in the direction toward the origin.

4. (a) Find a line through (3, 4) along which the rate of change of $f(x, y) = (169 - x^2 - y^2)^{1/2}$ at (3, 4) is equal to 0. (b) In what direction at $(\frac{5}{2}, 6)$ is $f$ *decreasing* most rapidly? Indicate the direction by specifying a unit vector with the required direction. (c) What is the rate of change of $f$ in the direction mentioned in (b)?

5. If $F(x, y, z) = 4x^2 + 9y^2 - 18z$, find $dF/ds$ at (3, 2, 1): (a) along the line $\dfrac{x-3}{2} = \dfrac{y-2}{3} = \dfrac{z-1}{1}$ in the direction of increasing $x$; (b) along the normal to the plane $4(x - 3) + (y - 2) = z - 1$ in the direction of increasing $z$; (c) in the direction of most rapid increase of $F$.

6. Find the component of grad $F$ in each of the directions indicated if $F(x, y, z) = x^2y - 3y^2z + 4z^2x - 6xyz$.
   (a) At (1, 2, 1) in the direction of the vector $2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.
   (b) At (1, 0, 2) in the direction of $(\mathbf{i} + \mathbf{j}) \times \text{grad } F$.
   (c) At (1, 0, 2) in the direction toward (5, 4, 4).
   (d) At (3, 4, 12) in the direction toward the origin.

## 19-8 Implicitly Defined Functions

In Example 3 and Exercises 2 and 9 of § 19-2 we saw how to find $\partial z/\partial x$ and $\partial z/\partial y$ in simple cases when $z$ is to be regarded as a function of $x$, $y$ defined implicitly by an equation in $x$, $y$, $z$. In this section we return to the subject of implicit functions, but now we are concerned with general formulas as well as with technique in particular problems.

If $F$ is a differentiable function of three variables, and if $f$ is a differentiable function of two variables such that

$$F[x, y, f(x, y)] = 0 \tag{1}$$

for all points in the interior of some region in the $xy$-plane, we can apply the chain rule to (1), with $x$, $y$, $z$ as variables of the first class in $F(x, y, z)$, and with $x$, $y$ as variables of the second class, where

$$x = x, \qquad y = y, \qquad z = f(x, y) \tag{2}$$

are the equations connecting the two classes. Then if $G(x, y) = F[x, y, f(x, y)]$, we see that

$$0 = \frac{\partial G}{\partial x} = \frac{\partial F}{\partial x}\cdot 1 + \frac{\partial F}{\partial y}\cdot 0 + \frac{\partial F}{\partial z}\frac{\partial f}{\partial x}, \tag{3}$$

with a similar equation expressing the fact that $0 = \partial G/\partial y$. Here $\partial F/\partial x = F_1(x, y, z)$, with $z = f(x, y)$, and so on. Then, if $F_3[x, y, f(x, y)]$

$\neq 0$, we can solve for $\partial z/\partial x$ in (3), obtaining

$$\frac{\partial f}{\partial x} = -\frac{F_1[x, y, f(x, y)]}{F_3[x, y, f(x, y)]}. \tag{4}$$

**Example 1:** To illustrate this general formula, suppose that

$$\left.\begin{aligned} f(x, y) &= \log \frac{1 - x^2 - y^2}{1 + x^2 + y^2}, \\ F(x, y, z) &= (1 + x^2 + y^2)e^z + x^2 + y^2 - 1. \end{aligned}\right\} \tag{5}$$

With these definitions of $f$ and $F$ it is clear that (1) holds true. In this case (3) is just what we get if we set

$$(1 + x^2 + y^2)e^z + x^2 + y^2 - 1 = 0$$

and differentiate with respect to $x$, treating $x$ and $y$ as independent and $z$ as dependent:

$$2xe^z + 2x + (1 + x^2 + y^2)e^z \frac{\partial z}{\partial x} = 0.$$

Solving, we obtain

$$\frac{\partial z}{\partial x} = -\frac{2x(e^z + 1)}{(1 + x^2 + y^2)e^z}.$$

If in this result we replace $z$ by $f(x, y)$ as defined in (5), we obtain

$$e^z = \frac{1 - x^2 - y^2}{1 + x^2 + y^2}, \qquad e^z + 1 = \frac{2}{1 + x^2 + y^2},$$

$$\frac{\partial z}{\partial x} = -\frac{4x}{(1 + x^2 + y^2)(1 - x^2 - y^2)}.$$

This agrees with what would be obtained by direct differentiation of $z = f(x, y)$. We leave verification to the student.

For emphasis on results we state the following theorem, which recapitulates the result (4) and the corresponding result with $y$ in place of $x$.

THEOREM 19-E. *Let $F$ be a differentiable function of $x$, $y$, $z$ and let $f$ be a differentiable function of $x$ and $y$. Suppose that $F(x, y, z) = 0$ and $\partial F/\partial z \neq 0$ at each point on the surface defined by $z = f(x, y)$, where the point $(x, y)$ varies over a certain region of the $xy$-plane. Then*

$$\frac{\partial f}{\partial x} = -\frac{F_1(x, y, z)}{F_3(x, y, z)}, \qquad \frac{\partial f}{\partial y} = -\frac{F_2(x, y, z)}{F_3(x, y, z)}, \tag{6}$$

*and for the normal to the surface we have*

$$\textit{direction of normal} = \frac{\partial F}{\partial x} : \frac{\partial F}{\partial y} : \frac{\partial F}{\partial z}, \tag{7}$$

*where it is understood in (6) and (7) that the partial derivatives of $F$ are evaluated at the point $(x, y, z)$ of the surface $z = f(x, y)$.*

*Proof.* The first formula in (6) is just the same as (4). The assumption that $F_3(x, y, z) \neq 0$ is needed in order to obtain (4) from (3). The second formula in (6) is obtained by the same argument, by focusing attention on $y$ instead of on $x$. To obtain (7) we start from the fact, explained in §19-2, that the direction of the normal is

$$\frac{\partial f}{\partial x} : \frac{\partial f}{\partial y} : -1.$$

But the vector with components $\partial F/\partial x$, $\partial F/\partial y$, $\partial F/\partial z$ is just $-\partial F/\partial z$ times the vector with components $\partial f/\partial x$, $\partial f/\partial y$, $-1$. Since the latter vector is normal to the surface $z = f(x, y)$ at $(x, y, z)$, so is the former vector, and this justifies (7).

We are acquainted with examples of surfaces defined by equations of the form $F(x, y, z) = 0$. Spheres and ellipsoids are illustrations. For instance, in the case of the ellipsoid

$$\frac{x^2}{36} + \frac{y^2}{25} + \frac{z^2}{16} = 1,$$

we can take

$$F(x, y, z) = \frac{x^2}{36} + \frac{y^2}{25} + \frac{z^2}{16} - 1. \tag{8}$$

Now, if we consider the locus defined by the equation $F(x, y, z) = 0$, it may be that *part* of this locus can be represented by an equation of the form $z = f(x, y)$. In the case of the foregoing ellipsoid we can represent its upper part by

$$z = 4\sqrt{1 - \frac{x^2}{36} - \frac{y^2}{25}} \tag{9}$$

and its lower part by

$$z = -4\sqrt{1 - \frac{x^2}{36} - \frac{y^2}{25}}. \tag{10}$$

Here either (9) or (10) could play the role of $z = f(x, y)$ in Theorem 19-E. The limitation on $(x, y)$ to make $f$ a differentiable function is expressed in either case by the inequality

$$\frac{x^2}{36} + \frac{y^2}{25} < 1. \tag{11}$$

With this condition in force we have $F(x, y, z) = 0$ and $\partial F/\partial z \neq 0$ on the part of the surface in question. Note here that $\partial F/\partial z = z/8 \neq 0$ is a result of (11) with either (9) or (10).

There are general theorems concerned with the question of when the equation $F(x, y, z) = 0$ can be solved (in a theoretical sense) for $z$ as a function of $x$ and $y$ in such a way that the conditions of Theorem 19-E will be satisfied. These are called *implicit function theorems*. They are beyond the scope of this book.

Implicit function situations can arise with different numbers of variables, and also with more than one equation. We illustrate as follows. Suppose two surfaces are defined by the equations

$$F(x, y, z) = 0, \qquad G(x, y, z) = 0,$$

respectively, and suppose that they intersect in such a way that part of the intersection is a curve which can be represented by expressing $y$ and $z$ as functions of $x$, say

$$y = f(x), \qquad z = g(x). \tag{12}$$

Assuming that all the functions mentioned are differentiable, let us see how to compute the direction of the tangent to the curve. We know from § 18-6 that this direction is

$$dx:dy:dz = 1:f'(x):g'(x). \tag{13}$$

Now, according to our assumptions,

$$F[x, f(x), g(x)] = 0, \qquad G[x, f(x), g(x)] = 0.$$

Using the chain rule, we have

$$\left.\begin{aligned} \frac{\partial F}{\partial x}\cdot 1 + \frac{\partial F}{\partial y} f'(x) + \frac{\partial F}{\partial z} g'(x) = 0, \\ \frac{\partial G}{\partial x}\cdot 1 + \frac{\partial G}{\partial y} f'(x) + \frac{\partial G}{\partial z} g'(x) = 0. \end{aligned}\right\} \tag{14}$$

We now regard these two equations as simultaneous linear equations to be solved for $f'(x)$ and $g'(x)$. In order to be able to solve uniquely we need to assume that

$$\begin{vmatrix} \dfrac{\partial F}{\partial y} & \dfrac{\partial F}{\partial z} \\ \dfrac{\partial G}{\partial y} & \dfrac{\partial G}{\partial z} \end{vmatrix} \neq 0 \tag{15}$$

when the partial derivatives are evaluated at points on the curve of intersection. Particular cases of this will be found in the exercises.

*Extremal Problems with Side Conditions*

The concept of an extremal problem with a side condition, as explained in § 3-11, occurs also with functions of several variables. The side conditions often lead to situations where it is advantageous to use implicitly defined functions.

**Example 2:** Find the maximum value of $G(x, y, z) = ax + by + cz$ on the surface of the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \tag{16}$$

assuming that $a$, $b$, $c$ are all positive. Take for granted that a maximum exists and that when it occurs, $x$, $y$, and $z$ are all different from zero. It is evident that none of them can be negative when the maximum is attained.

Under the given conditions we think of $z$ as a function of $x$, $y$, say $z = f(x, y)$, such that (16) holds, and we look for a critical point of the function $ax + by + cf(x, y)$, with $x$, $y$ as independent variables. Then

$$\frac{\partial}{\partial x}(ax + by + cz) = a + c\frac{\partial z}{\partial x}, \qquad \frac{\partial}{\partial y}(ax + by + cz) = b + c\frac{\partial z}{\partial y},$$

so that we have the equations

$$a + c\frac{\partial z}{\partial x} = 0, \qquad b + c\frac{\partial z}{\partial y} = 0.$$

But also, from (16),

$$\frac{2x}{a^2} + \frac{2z}{c^2}\frac{\partial z}{\partial x} = 0, \qquad \frac{2y}{b^2} + \frac{2z}{c^2}\frac{\partial z}{\partial y} = 0.$$

We now eliminate the derivatives $\partial z/\partial x$, $\partial z/\partial y$, obtaining

$$\frac{x}{a^2} + \frac{z}{c^2}\left(-\frac{a}{c}\right) = 0, \quad \text{or} \quad x = \frac{a^3}{c^3}z,$$

and likewise, because of symmetry,

$$y = \frac{b^3}{c^3}z.$$

Going back to (16) and eliminating $x$ and $y$, we have

$$\frac{a^4}{c^6}z^2 + \frac{b^4}{c^6}z^2 + \frac{z^2}{c^2} = 1, \quad \text{or} \quad z^2 = \frac{c^6}{a^4 + b^4 + c^4}.$$

Thus, taking the positive square root, we find

$$z = \frac{c^3}{(a^4 + b^4 + c^4)^{1/2}}.$$

The values of $x$ and $y$ are similar, with $a^3$ and $b^3$ replacing $c^3$. Thus the required maximum value of $ax + by + cz$ is

$$\frac{a^4}{(a^4 + b^4 + c^4)^{1/2}} + \frac{b^4}{(a^4 + b^4 + c^4)^{1/2}} + \frac{c^4}{(a^4 + b^4 + c^4)^{1/2}} = (a^4 + b^4 + c^4)^{1/2}.$$

## EXERCISES

**1.** In each case a pair of equations is given. It is assumed that $u$ and $v$ are differentiable functions of $x$ and $y$ such that the equations are satisfied. Find $\partial u/\partial x$, $\partial u/\partial y$, $\partial v/\partial x$, and $\partial v/\partial y$, assuming that the appropriate second order determinants are not zero.

(a) $u + v - x^2 = 0,$
$u^2 - v^2 - y = 0.$

(b) $u^2 - v + x = 0,$
$u + v^2 - y = 0.$

(c) $e^u \cos v - x = 0,$
$e^u \sin v - y = 0.$

(d) $u^2 - v^2 - xy = 0,$
$uv + x^2 - y^2 = 0.$

**2.** Work out the equations corresponding to (14) for the case of the curve of intersection of the sphere and ellipsoid

$$x^2 + y^2 + z^2 = 25, \qquad \frac{x^2}{36} + \frac{y^2}{25} + \frac{z^2}{16} = 1.$$

and solve for $dy/dx$ and $dz/dx$. Compare the result with what you get if you first solve for $y^2$ and $z^2$ in terms of $x$ and then differentiate. What does the condition (15) become in this case?

**3.** (a) If $F$ is a differentiable function of $x$ and $y$, if $C = F(x_0, y_0)$, and if $F_2(x_0, y_0) \neq 0$, how is the slope at $(x_0, y_0)$ of the level curve $F(x, y) = C$ expressed in terms of partial derivatives of $F$? (b) Use the result in (a) to show that the gradient of $F$ at $(x_0, y_0)$ is perpendicular to the level curve through the point.

**4.** Under the conditions stated in Theorem 19-E show that the gradient of $F$ at a point of the surface $z = f(x, y)$ is perpendicular to the surface at that point. Note that the surface is part of a level surface of the function $F$.

**5.** If $F$, $G$, and $H$ are differentiable functions of five variables, and if $f$, $g$, and $h$ are differentiable functions of two variables such that

$$F[x, y, f(x, y), g(x, y), h(x, y)] = 0$$

at all points $(x, y)$ in the interior of a certain region in the $xy$-plane, with similar equations for $G$ and $H$, show how to find the first partial derivatives of $f$, $g$, $h$ with respect to $x$, $y$ if a certain determinant is not zero. Write out the formula for $\partial f/\partial y$ explicitly.

**6.** Find the maximum value of $ax + by + cz$ subject to the condition $x^4 + y^4 + z^4 = 1$, assuming that $a$, $b$, $c$ are all positive, and taking for granted that the maximum occurs when $x$, $y$, $z$ are all different from zero.

**7.** Find the minimum of $ax + by + cz$ subject to the conditions that $ax^{-1} + by^{-1} + cz^{-1} = 1$, and that $x$, $y$, $z$ all be positive. Assume $a$, $b$, $c$ all positive and take for granted the existence of the minimum.

**8.** Find the maximum value of $\dfrac{xyz}{8x + 27y + 64z}$ subject to the conditions that $xyz = 64$ and $x$, $y$, $z$ are all positive. Take for granted that the maximum exists.

**9.** If $z$ is a function of $x$, $y$ such that $z^3 + 3xy - 3y = 0$, find $\dfrac{\partial^2 z}{\partial x^2}$, $\dfrac{\partial^2 z}{\partial x\, \partial y}$ and $\dfrac{\partial^2 z}{\partial y^2}$ if $z^2 + x \neq 0$. Show that $\dfrac{\partial^2 z}{\partial x^2} = -x\dfrac{\partial^2 z}{\partial y^2}$.

**10.** With $f$ and $F$ as in Theorem 19-E, but having continuous second partial derivatives, show that

$$\frac{\partial^2 f}{\partial x^2} = \frac{F_3^2 F_{11} - 2F_1F_3F_{13} + F_1^2F_{33}}{-F_3^3},$$

and derive an analogous formula for $\partial^2 f/\partial x\, \partial y$. Here it is understood that $F_3$, $F_{11}$, and so on are evaluated at a point $(x, y, z)$ on the surface $z = f(x, y)$.

**11.** Find, both by analytic geometry and by calculus, the points on the sphere $x^2 + y^2 + z^2 - 4x - 6y - 10z + 14 = 0$ which are, respectively, farthest from and nearest to the point (3, 5, 4).

**12.** A sheet metal container is to be made of a right circular cylinder with equal right circular conical caps on the ends. Show that, for a fixed volume, the total surface area is least when the length of the cylinder is the same as the altitude of each cone and the diameter of the cylinder is $\sqrt{5}$ times the length.

# CHAPTER XX

# MULTIPLE INTEGRALS

## 20-1 Double Integrals

A double integral of a function of two variables is the two-dimensional analogue of a definite integral of a function of one variable. It is now convenient to call this latter type of integral a *single integral,* to contrast with a *double integral.*

The value of the single integral $\int_a^b f(x)\,dx$ is determined by the function $f$ and the interval $[a, b]$. For the case of a double integral of a function of $x$ and $y$, the role of the interval $[a, b]$ is taken by a region $R$ in the $xy$-plane, and the double integral of $f(x, y)$ over $R$ is denoted by

$$\iint_R f(x, y)\,dx\,dy \quad \text{or} \quad \iint_R f(x, y)\,dA. \tag{1}$$

The reason for the $dA$ notation will be explained presently.

A double integral is defined as a limit of certain sums, in much the same way that the single integral $\int_a^b f(x)\,dx$ was defined in § 6-1. Moreover, double integrals have applications similar in nature to the applications of single integrals. For students at this stage of the study of calculus there are essentially three aspects of the study of double integrals: (1) the definition of the integral as a limit of sums; (2) the formulation of various geometrical and physical magnitudes as definite integrals; (3) the methods of computing values of double integrals.

We begin with the definition in the simplest case. Suppose that $R$ is a rectangular region consisting of all points $(x, y)$ such that $a \leq x \leq b$

and $c \leq y \leq d$, where $a$, $b$, $c$, $d$ are numbers such that $a < b$ and $c < d$. Some or all of these numbers may be negative. Let $f$ be a function of $x$ and $y$ which is defined at each point of $R$. The function might perhaps be defined at some points not in $R$, but we ignore this and regard $R$ as the domain of definition of $f$. Ordinarily we consider situations in which $f$ is continuous at each point of $R$, but it would do no harm to have certain types of discontinuous behavior of $f$. However, we shall not attempt to describe precisely what might be permissible in this respect, and we shall for the present assume that $f$ is continuous at each point of $R$.

*Definition of a Double Integral*

Now let $R$ be divided into a number of smaller rectangles in the following manner: Choose numbers $x_0, x_1, \cdots, x_m$ and $y_0, y_1, \cdots, y_n$ so that

$$a = x_0 < x_1 < \cdots < x_m = b, \qquad c = y_0 < y_1 < \cdots < y_n = d,$$

and draw the various lines $x = x_j$, $y = y_k$, so that $R$ is divided into $mn$ rectangles, as shown in Fig. 20-1. We write $\Delta x_j = x_j - x_{j-1}$, $\Delta y_k = y_k - y_{k-1}$. Let $R_{jk}$ be the rectangle with sides $x = x_{j-1}$, $x = x_j$, $y = y_{k-1}$, $y = y_k$ (shaded in Fig. 20-1). In $R_{jk}$ choose any point $x = s_j$, $y = t_k$, and form the product $f(s_j, t_k)\,\Delta x_j\,\Delta y_k$, which is the value of $f$ at $(s_j, t_k)$ multiplied by the area of $R_{jk}$. Then form the sum of all these products; this sum can be written in the form

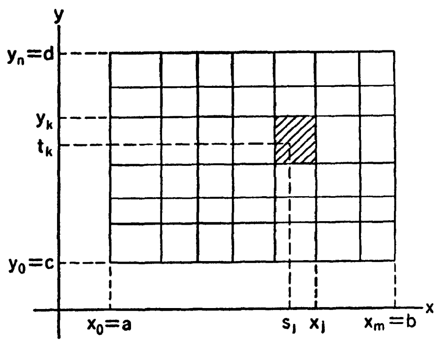



**Fig. 20-1**

$$\sum_{j=1}^{m} \sum_{k=1}^{n} f(s_j, t_k)\,\Delta x_j\,\Delta y_k \tag{2}$$

by using a summation symbol notation. Now consider what happens as the maximum of all the numbers $\Delta x_1, \cdots, \Delta x_m, \Delta y_1, \cdots, \Delta y_n$ is made to approach zero. (This will, of course, force $m$ and $n$ to increase indefinitely.) It turns out that the sums (2) approach a definite limit, and this limit is called the value of the double integral of $f$ over the region $R$. The value of the integral is indicated by either of the notations in (1). The meaning of the sums approaching their limit is that the absolute value

$$\left|\iint_R f(x, y)\,dA - \sum_{j=1}^{m} \sum_{k=1}^{n} f(s_j, t_k)\,\Delta x_j\,\Delta y_k\right| \tag{3}$$

can be made as small as we please, simply by making the greatest of the $\Delta x_j$'s and $\Delta y_k$'s sufficiently small. Apart from this condition on the $\Delta x_j$'s and $\Delta y_k$'s it does not matter how the points $x_j$ and $y_k$ are spaced, nor how the points $(s_j, t_k)$ are chosen in $R_{jk}$. The $dA$ notation in (3) is suggested by using $\Delta A_{jk}$ instead of $\Delta x_j\, \Delta y_k$ to denote the area of $R_{jk}$; the notation is purely conventional, following historical tradition, for $dA$ is not the differential of a function.

The fact that the sums (2) do approach a limit can be proved as a consequence of the continuity of the function $f$. This proof is given in more advanced textbooks on calculus.

It is necessary to define double integrals over regions of somewhat arbitrary shape, as well as over rectangles. The procedure is much as before, but there are some differences, owing to the fact that if the boundaries of the region are curved, the region cannot be exactly filled out by small rectangles.

Suppose now that $R$ is a region which is bounded by one or more closed curves. We suppose that the curves are composed of a finite number of simple arcs, each defined either by specifying $y$ as a function of $x$ on some interval of the $x$-axis, or the corresponding situation with the roles of $x$ and $y$ reversed; these functions shall be continuous. The simplest typical case would be where there is just one closed circuit (e.g., a triangle, a polygon, an ellipse, or a semicircle and a diameter) forming the boundary of $R$. But $R$ might also be such a thing as a circular region with a square hole cut out of it. Let two sets of lines be drawn, one set parallel to the $x$-axis, the other set parallel to the $y$-axis. The spacing of the lines need not be regular, but the lines should be close enough together so that the rectangles formed (see Fig. 20-2) are small in comparison with the size of $R$. The network of lines forms what we call a *rectangular partition;* an individual rectangle in the partition is called a *cell.* Some of the cells will lie entirely in the region $R$; other cells will contain points not in $R$. For our purposes we retain only those cells which do not in any way extend outside of $R$. We then number the retained cells in some arbitrary order. If there are $N$ cells, let their areas be $\Delta A_1, \cdots, \Delta A_N$. Now suppose that $f$ is a function which is continuous at each point of $R$. Choose any point $(x_i, y_i)$ in the $i$th cell, and form the sum

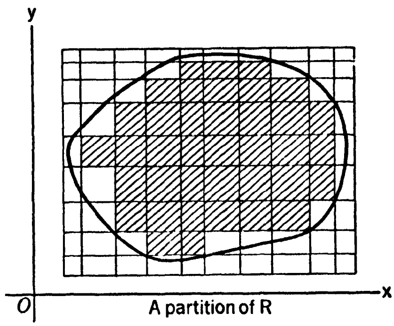



**Fig. 20-2**

$$\sum_{i=1}^{N} f(x_i, y_i)\, \Delta A_i. \tag{4}$$

We then take the limit of this sum in the same manner as before, and define the limit as the value of the double integral of $f$ over $R$.

In forming the sums (4), no trouble is caused by the small, irregularly shaped fragments of the region $R$ which are not covered by any of the retained cells. The total area of the part of $R$ which is in these fragments approaches zero in the limiting process.

It is clear that if $f$ is constant in value, say $f(x, y) = c$, then

$$\iint_R f(x, y)\, dA = cA,$$

where $A$ is the area of the region $R$.

Now we turn to some applications of the double integral concept. Our discussion will help to make plausible the fact that the sums (2) and (4) do indeed approach limits. Furthermore, we shall in the process learn some of the principal uses of double integrals.

*Volume Under a Surface*

Suppose that $f$ is a continuous function with values which are never negative at points of the region $R$. Then the surface $z = f(x, y)$ forms a canopy over the region $R$, and the volume $V$ of the space directly under this canopy and directly above $R$ is given by

$$V = \iint_R f(x, y)\, dA. \tag{5}$$

That this is so is seen by a discussion of the definition of volume for a solid figure with curved bounding surfaces, quite analogous to the discussion of area between a curve $y = g(x)$ and the $x$-axis, as given in connection with Fig. 2-24 and Fig. 2-25 in § 2-6, and in the related discussion in § 6-1. One of the terms $f(x_i, y_i)\, \Delta A_i$ in the sum (4) is the volume of a rectangular parallelepiped of height $f(x_i, y_i)$ and base area $\Delta A_i$ (see Fig. 20-3). If we choose $(x_i, y_i)$ so that $f(x_i, y_i)$ is the smallest value of $f$ in the $i$th cell, the product $f(x_i, y_i)\, \Delta A_i$ then represents the volume of a parallelepiped which in general does not fill up the column of space under the surface and on the $i$th cell. On the other hand, if $(x_i, y_i)$ is chosen so that $f(x_i, y_i)$ is the largest value of $f$ in the $i$th cell, we get the volume of a parallelepiped which in general more than fills up the

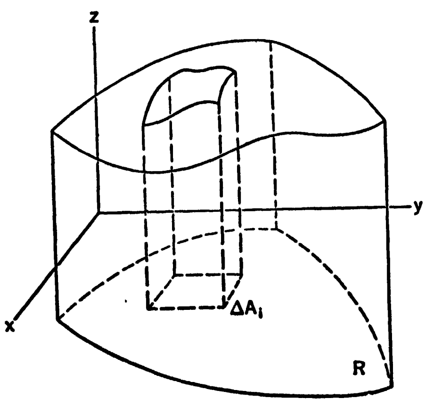



**Fig. 20-3**

column in question. Hence it is seen that (5) is correct essentially as a *definition* of the total volume of the solid with base $R$, with top surface the surface $z = f(x, y)$, and bounded laterally by the cylindrical surface formed by drawing lines parallel to the $z$-axis through the boundary of $R$.

### *Laminas of Variable Density*

When we think of mass as spread over a plane region, as in the case of a circular disk covered with a thin layer of gold leaf of varying thickness, we form the concept of an *areal density*, or mass per unit area, as follows. The density at the point $(x, y)$ is the value at $(x, y)$ of a function, called the density function. If we denote the density function by $\sigma$, then its principal property is that

$$M = \iint_R \sigma(x, y)\, dA \tag{6}$$

is the total mass which is spread over the region $R$. This holds true by definition when $R$ is any part of the total region over which mass is spread. It is usually assumed that $\sigma$ is a continuous function whose values are never negative, and that points where $\sigma(x, y) = 0$ are exceptional in the sense that no region, however small, has total mass zero. This does permit $\sigma(x, y)$ to be zero at isolated points or along certain curves.

To get the density directly from the mass, we can think of $\sigma(x, y)$ as the limit of $\Delta M/\Delta A$, where $\Delta A$ is the area and $\Delta M$ is the mass for a small region containing $(x, y)$. Here the limit is taken as the region closes down on $(x, y)$; i.e., as the region shrinks in such a way that the maximum distance from $(x, y)$ to any point of the region approaches zero.

A plane region which carries a spread of mass is called a *lamina*. Later we shall consider curved laminas also, i.e., pieces of a curved surface on which there is a spread of mass.

To locate the center of mass of a lamina we use the ideas developed in §11-5. We have to define the total moment of the mass of the lamina about each coordinate axis. We shall see that the total moment about the $y$-axis is properly defined to be

$$\iint_R x\sigma(x, y)\, dA, \tag{7}$$

and hence, if $(\bar{x}, \bar{y})$ is the center of mass, then $\bar{x}$ is found from the equation

$$M\bar{x} = \iint_R x\sigma(x, y)\, dA, \tag{8}$$

where $M$ is given by (6). To justify (7) we form an auxiliary system of particles by forming a rectangular partition of the lamina as in Fig. 20-2, and thinking of the mass of the $i$th cell as being concentrated at some point in it. Now, the mass of the $i$th cell is its area times a suitable mean value of the density. Since the density is assumed to be continuous, the mass of the cell is

$$\Delta M_i = \sigma(x_i, y_i)\, \Delta A_i,$$

where $(x_i, y_i)$ is so chosen in the cell that $\sigma(x_i, y_i)$ is the appropriate mean value of the density. We now think of the total mass of the cell as being concentrated at $(x_i, y_i)$. The total moment about the $y$-axis of the auxiliary system of particles is then

$$\sum_{i=1}^{N} x_i \sigma(x_i, y_i)\, \Delta A_i.$$

The limit of this sum, when the maximum cell dimension approaches zero, is the double integral (7), and so our definition is justified.

There is, of course, a formula similar to (8) for finding $\bar{y}$. Examples of the actual calculation of $\bar{x}$ and $\bar{y}$ will be given in a later section.

*Moments of Inertia*

For basic ideas about moments of inertia we refer back to § 6-10, which should be re-read by the student at this point. If we are to find the moment of inertia of a lamina about an axis $L$, let $D(x, y)$ be the perpendicular distance from the point $(x, y)$ to the axis $L$, which *may* be in the same plane as the lamina, but *need not* be. If we use the same auxiliary system of particles as in the discussion of (7), the moment of inertia of the system about the axis $L$ is

$$\sum_{i=1}^{N} [D(x_i, y_i)]^2\, \Delta M_i = \sum_{i=1}^{N} [D(x_i, y_i)]^2 \sigma(x_i, y_i)\, \Delta A_i,$$

and hence we are led to the integral

$$I = \iint_R [D(x, y)]^2 \sigma(x, y)\, dA \qquad (9)$$

as the proper formulation of the moment of inertia of the lamina. We observe that, if $L$ is the $z$-axis (perpendicular to the $xy$-plane at the origin), then $[D(x, y)]^2 = x^2 + y^2$, whereas if $L$ is the $y$-axis, then $[D(x, y)]^2 = x^2$.

## 20-2 Iterated Integrals

In this section we shall show how to calculate the value of a double integral by performing two successive single integrations. We begin with the case in which the double integral represents a volume, as in § 20-1, (5), and we revert to § 6-7, where it was explained how to find the volume of a solid by slicing it with planes perpendicular to the $x$-axis. The formula for a volume, as obtained in § 6-7, was

$$V = \int_a^b A(x)\,dx, \tag{1}$$

where $A(x)$ is the area of the cross section of the solid made by the plane determined by an arbitrary value of $x$, and the complete volume is obtained as $x$ varies from $a$ to $b$, where $a < b$.

We shall apply (1) to the volume $V$ given by the double integral

$$V = \iint_R f(x, y)\,dx\,dy, \tag{2}$$

where $f$ is a continuous function with values which are positive at points of $R$. We assume that $R$ is a region of the type shown in Fig. 20-4. That is, let $g_1$ and $g_2$ be two continuous functions of $x$ defined on the same interval $[a, b]$ and such that $g_1(x) < g_2(x)$ if $a < x < b$. The region $R$ consists of all points $(x, y)$ such that $a \leq x \leq b$ and $g_1(x) \leq y \leq g_2(x)$.

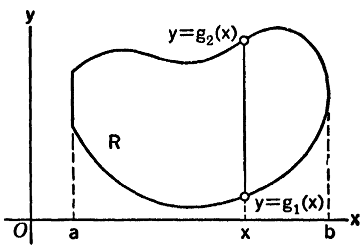



**Fig. 20-4**

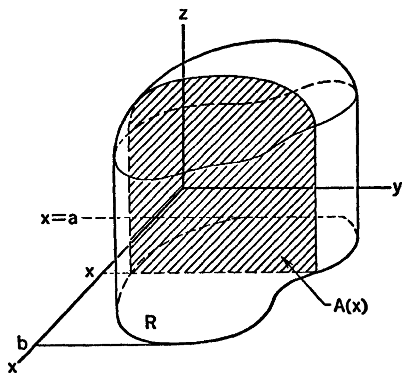



**Fig. 20-5**

The solid whose volume we wish to find is shown in Fig. 20-5. The diagram also shows the typical cross section whose area $A(x)$ enters in formula (1). Evidently $A(x)$ can be calculated by an integration of $z = f(x, y)$ with respect to $y$, keeping $x$ fixed. The formula is

$$A(x) = \int_{g_1(x)}^{g_2(x)} f(x, y)\,dy. \tag{3}$$

Hence, equating $V$ in (1) and (2), we have

$$\iint_R f(x, y)\,dx\,dy = \int_a^b \left[\int_{g_1(x)}^{g_2(x)} f(x, y)\,dy\right] dx. \tag{4}$$

Here, on the right, the inside integration with respect to $y$ is performed first, yielding a result which is a function of $x$. Then the $x$ integration is performed. It is customary to omit the brackets around the inside integral and relocate the $dx$, so that (4) becomes

$$\iint_R f(x, y)\, dx\, dy = \int_a^b dx \int_{g_1(x)}^{g_2(x)} f(x, y)\, dy. \tag{5}$$

The expression on the right is called an *iterated* integral, or a *repeated* integral. This formula gives us a means of calculating the value of the double integral.

**Example 1:** Find the volume under the plane $12x + 10y + 15z = 60$ and above the triangle in the $xy$-plane bounded by the lines $y = 0$, $x = 2$, $5y = 9x$. (The volume in question is that of the solid $OABCD$ in Fig. 20-6.)

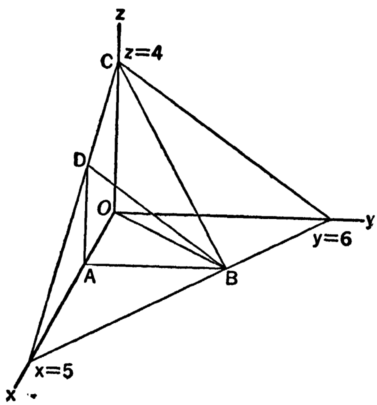



Fig. 20-6

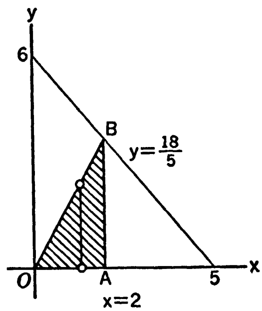



Fig. 20-7

The required volume is

$$V = \iint_R \frac{60 - 12x - 10y}{15}\, dx\, dy,$$

where $R$ is the triangle referred to. From Fig. 20-7 we see that $g_1(x) = 0$ and $g_2(x) = 9x/5$ in this case. Hence, by (5),

$$V = \int_0^2 dx \int_0^{9x/5} \frac{60 - 12x - 10y}{15}\, dy.$$

The first integration is as follows

$$\frac{1}{15}\int_0^{9x/5} (60 - 12x - 10y)\, dy = \frac{1}{15}\Big[60y - 12xy - 5y^2\Big]_0^{9x/5}$$

$$= \frac{1}{15}\left[108x - \frac{108}{5}x^2 - \frac{81}{5}x^2\right]$$

$$= \frac{36}{5}x - \frac{63}{25}x^2.$$

Then
$$V = \int_0^2 \left(\frac{36}{5}x - \frac{63}{25}x^2\right) dx$$
$$= \left[\frac{18}{5}x^2 - \frac{21}{25}x^3\right]_0^2 = \frac{72}{5} - \frac{168}{25} = \frac{192}{25}.$$

This result can be checked by considering the volume as that of a pyramid with trapezoidal base $OADC$ and altitude $AB$.

Our derivation of (5) was based on the identification of (1) and (2) as expressions for a certain volume. What we want now is to establish (5) as a general formula, valid even when the values of $f$ need not all be positive. To establish this we need an argument that is independent of the interpretation of the double integral as a volume. A fully accurate analytical proof of (5), for the case in which $f$ is an arbitrary continuous function, is rather long, and we shall not give it here. (A proof is given in §16.61 of the author's *Advanced Calculus*, Ginn & Company, 1955.) The gist of the proof, with the suppression of some details, can be expressed in the following way. In Fig. 20-4 let $c$ be the minimum of $g_1(x)$ on $[a, b]$, and likewise let $d$ be the maximum of $g_2(x)$ on $[a, b]$. Divide the interval $[a, b]$ into $m$ equal parts by points $x_0, x_1, \cdots, x_m$, and divide $[c, d]$ into $n$ equal parts by points $y_0, y_1, \cdots, y_n$. Now consider the cells into which the rectangle with sides $x = a, x = b, y = c, y = d$ is divided by the various lines $x = x_j$, $y = y_k$. Form the sum

$$\sum_j \sum_k f(x_j, y_k)\, \Delta x_j\, \Delta y_k, \tag{6}$$

where $\Delta x_j = x_j - x_{j-1}$, $\Delta y_k = y_k - y_{k-1}$, and the sum includes all terms for which the point $(x_j, y_k)$ is in $R$. As $m$ and $n$ become very large, the sum (6) approaches the double integral $\iint_R f(x, y)\, dx\, dy$ as limit. On the other hand, suppose that we keep $m$ fixed and consider what happens to the sum (6) as $n \to \infty$. For a fixed $j$, the definition of a single integral shows that

$$\lim_{n\to\infty} \sum_k f(x_j, y_k)\, \Delta y_k = \int_{g_1(x_j)}^{g_2(x_j)} f(x_j, y)\, dy. \tag{7}$$

The limits of integration here are what they are shown to be because the sum with respect to $k$ includes only cells for which $(x_j, y_k)$ is in $R$ (see Fig. 20-8). From (7) we conclude that when $n$ is large, the sum (6) is approximately equal to

$$\sum_j \left(\int_{g_1(x_j)}^{g_2(x_j)} f(x_j, y)\, dy\right) \Delta x_j.$$

When $m \to \infty$, the limit of this last sum is

$$\int_a^b \left[\int_{g_1(x)}^{g_2(x)} f(x, y)\, dy\right] dx. \tag{8}$$

In a more detailed argument it could be shown that in this case the limit of the sum (6) when $m$ and $n$ simultaneously become infinite is the same as the limit obtained by letting first $n$ and then $m$ become infinite. When these details are attended to, we get a proof of (5).

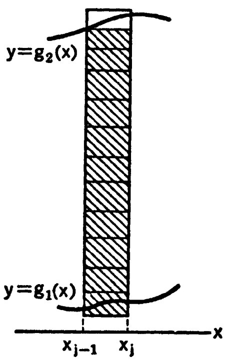



**Fig. 20-8**

The main new thing to be learned, as a matter of procedure in evaluating double integrals, is how to determine the limits of integration in the iterated integral. The student should study Fig. 20-4 and its relation to formula (5) until he thoroughly understands the method of putting the limits of integration on the integrals from the information provided by the diagram. Observe that the limits of integration are not affected by the function $f(x, y)$ which is being integrated.

There is, of course, an exactly analogous way of expressing the value of a double integral as an iterated integral first with respect to $x$ and then with respect to $y$. We illustrate with an example.

**Example 2:** Compute the moment about the $y$-axis of a homogeneous lamina of unit density, if the lamina occupies the smaller region $R$ cut from the circle $x^2 + y^2 = 4$ by the line $x + y = 2$ (see Fig. 20-9).

The required moment is given by the double integral

$$\iint_R x\,dA.$$

For a typical $y$ we see that $(x, y)$ is in $R$ if $2 - y \le x \le \sqrt{4 - y^2}$, for $x = 2 - y$ is the equation of the line and $x = \sqrt{4 - y^2}$ is the equation of the relevant part of the circle. To get all of $R$ we have to consider all $y$'s such that $0 \le y \le 2$. Hence

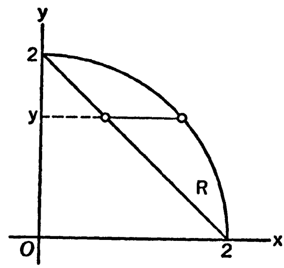



**Fig. 20-9**

$$\iint_R x\,dA = \int_0^2 dy \int_{2-y}^{\sqrt{4-y^2}} x\,dx.$$

Now
$$\int_{2-y}^{\sqrt{4-y^2}} x\,dx = \frac{x^2}{2}\Bigg|_{2-y}^{\sqrt{4-y^2}}$$
$$= \frac{4 - y^2}{2} - \frac{(2 - y)^2}{2} = 2y - y^2.$$

Then
$$\iint_R x\,dA = \int_0^2 (2y - y^2)\,dy = \left(y^2 - \frac{y^3}{3}\right)\Bigg|_0^2$$
$$= 4 - \frac{8}{3} = \frac{4}{3}.$$

## EXERCISES

**1.** Find the value of $\iint\limits_R f(x, y)\, dA$ in each of two ways for the functions and regions described.

(a) $f(x, y) = 2xy - x^2$; $R$ the triangle bounded by $x = -1$, $y = -1$, $4x + 3y = 5$.

(b) $f(x, y) = 2x - y$; $R$ the fourth quadrant portion of the interior of the circle $x^2 + y^2 = 9$.

(c) $f(x, y) = xy$; $R$ the region bounded by the parabola $y^2 = 4x$ and the line $2x - y = 4$. Note that when the first integration is with respect to $y$ the region should be broken into two parts, corresponding to $0 \le x \le 1$ and $1 \le x \le 4$.

**2.** In each case set up a double integral whose value is the volume described. Express the double integral in two ways as an iterated integral, and carry out the integration in one of the two orders.

(a) The volume of the tetrahedron cut from the first octant by the plane $4x + 3y + 2z = 12$.

(b) The volume of the tetrahedron with plane faces $y = 0$, $z = 0$, $x + y = 5$, $12x = 8y + 15z$.

(c) The volume of the tetrahedron with vertices $(0, 0, 0)$, $(0, 3, 0)$, $(1, 2, 0)$, $(0, 3, 4)$.

(d) The volume cut from the region inside the cylinder $x^2 + z^2 = b^2$ by the planes $y = 0$, $x = y$, $z = 0$.

(e) The volume enclosed between the paraboloid $b^2 z = a(b^2 - x^2 - y^2)$ and the $xy$-plane.

**3.** Follow the same directions as in Exercise 2.

(a) The volume under the plane $z = 2y$ and above the first quadrant area bounded by $y = 0$, $x = 3$, $x^2 + y^2 = 36$.

(b) The volume under the plane $z = x + y$ and above the area cut from the first quadrant by the ellipse $4x^2 + 9y^2 = 36$.

(c) The volume under the cylinder $y = z^2$ and above the area in the $xy$-plane bounded by $y = 0$ and $x^2 + 9y = 9$.

(d) The volume in the first octant bounded by the cylinder $x^2 = 4 - z$ and the planes $x = 0$, $y = 0$, $z = 0$, $4x + 3y = 12$.

(e) The volume in the first octant bounded by the parabolic cylinders $z = 9 - x^2$, $x = 3 - y^2$, $y = 0$, $x = 0$.

**4.** Find the volume enclosed by the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$.

**5.** In each case a lamina of variable density is described. Find its mass (using $c$ for the constant of proportionality) and locate the center of mass. All literal constants are assumed to be positive.

(a) Triangular lamina with vertices $(0, 0)$, $(a, 0)$, $(0, a)$; density proportional to the square of the distance from $(0, 0)$.

(b) Square lamina with diagonally opposite corners $(0, 0)$, $(b, b)$; density proportional to the square of the distance from $(0, 0)$.

(c) The lamina of (a), but with density proportional to the distance from the $y$-axis.
(d) Triangular lamina with vertices at (0, 0), $(a, 0)$, $(a, b)$; density proportional to distance from the side $x = a$.
(e) Lamina in the first quadrant, bounded by $bx^2 = a^2y$, $x = 0$, $y = b$; density $\sigma = cx$.
(f) The lamina of (e), but with $\sigma = c(b - y)$.

**6.** Locate the centroids of the following plane regions, using double integrals. All literal constants are assumed to be positive.
(a) The triangle with vertices (0, 0), $(a, 0)$, $(b, c)$, where $a > b$.
(b) The semicircular region $x^2 + y^2 \leq b^2$, $y \geq 0$.
(c) The region described by $b^2x^2 + a^2y^2 \leq a^2b^2$, $x \geq 0$, $y \geq 0$.
(d) The first quadrant region bounded by $by^2 = a^2x$, $x = b$, $y = 0$.
(e) The region bounded by $bx^2 = a^2y$ and $ay = bx$.

**7.** Locate the centroids of the following plane regions, using double integrals.
(a) The region in the first quadrant between $x = 0$, $x = 1$ and between $y = x - x^2$, $y^2 = 4x$.
(b) The region bounded by the two parabolas $y = x^2 + x$, $y = 2x^2 - 2$.
(c) The region defined by $\sqrt{y} \leq x \leq 2 - y$, $0 \leq y \leq 1$.

**8.** Each region is regarded as a lamina of unit density. Find the moments of inertia about the axes indicated. All literal constants are positive.
(a) The triangular region bounded by $y = 0$, $x = H$, $Hy = Bx$; axes $y = 0$, $x = 0$, and $x = H$.
(b) The first quadrant region bounded by $x = 0$, $y = H$, $B^2y = Hx^2$; axes $y = 0$, $x = 0$, $y = H$.
(c) The rectangular region bounded by $x = 0$, $x = 2a$, $y = -b$, $y = b$; axes $x = 0$, $y = 0$.
(d) The region bounded by $y^2 = 2ax$, $x = 2a$; axes $x = 0$, $y = 0$, $x = 2a$.

**9.** What volume is represented by the iterated integral

$$\int_0^a dx \int_0^x \sqrt{a^2 - x^2}\, dy?$$

Draw a figure showing the volume in question. What would be the iterated integral if the order of integration were reversed?

**10.** Calculate the value of $\iint_R x\, dA$ if $R$ is the part of the first quadrant between the circles $x^2 + y^2 = a^2$, $x^2 + y^2 = b^2$, where $0 < a < b$. Work the problem in two ways, corresponding to the two possible orders of integration.

## 20-3 Iterated Integrals in Polar Coordinates

Sometimes the calculation of a double integral is greatly simplified by use of polar coordinates. We shall explain the theory of expressing a double integral as an iterated integral in polar coordinates. Let the double integral be $\iint_R f(x, y)\, dA$, and suppose that $f(x, y)$ is transformed into $F(r, \theta)$ when we set $x = r \cos \theta$, $y = r \sin \theta$. For example, if $f(x, y) = xy^2$, this becomes $r \cos \theta\, (r \sin \theta)^2 = r^3 \cos \theta \sin^2 \theta$, which is $F(r, \theta)$.

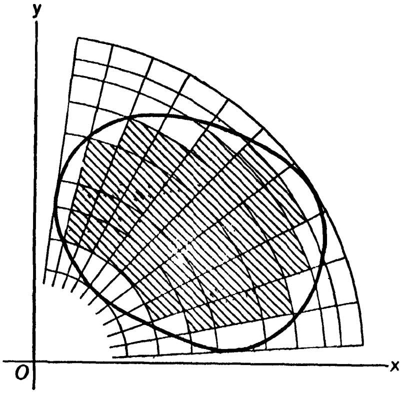



Fig. 20-10

Now let us form a polar coordinate partition of the plane by a series of circles with center at the origin and a series of rays emanating from $O$ (see Fig. 20-10). This partition of the plane forms cells which are rather like rectangles. We now proceed with a process much like that described in connection with Fig. 20-2. We select those cells which belong completely to the region $R$ (these cells are shaded in Fig. 20-10) and number them consecutively in some order. Suppose there are $N$ such cells. If $\Delta A_k$ is the area of the $k$th cell, and if $(x_k, y_k)$ is any point of the cell, it seems reasonable to suppose that

$$\lim \sum_{k=1}^{N} f(x_k, y_k)\, \Delta A_k = \iint_R f(x, y)\, dA, \tag{1}$$

the limit being taken in the sense that the partition is made finer and the cell size approaches zero. The truth of (1) is plausible if we think of the interpretation of the double integral as a volume; it is also plausible in the case when $f(x, y)$ is a density function and the integral is interpreted as the total mass of a lamina. We shall not attempt a formal proof of (1), but we shall use the result as basic in our argument. The whole subject can be treated rigorously in the theory of transformation of multiple integrals—a subject which is dealt with in books on advanced calculus.

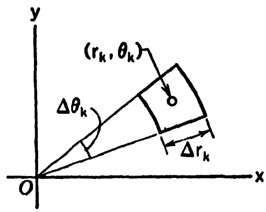



Fig. 20-11

The next step is to express $f(x_k, y_k)\, \Delta A_k$ in terms of polar coordinates. Consider the $k$th cell, as shown in Fig. 20-11. Let the polar coordinates of $(x_k, y_k)$ be $(r_k, \theta_k)$, and let us choose this point in the special position midway between the two circular arcs and midway between the two circular rays. Then the two circular arcs have radii

$$r = r_k - \tfrac{1}{2}\Delta r_k, \qquad r = r_k + \tfrac{1}{2}\Delta r_k.$$

The area of the cell is easily worked out by elementary geometry, starting from the formula for the area of a circular sector in terms of its radius and angular opening. The formula, which should be derived by the student, is

$$\Delta A_k = r_k \, \Delta r_k \, \Delta\theta_k. \tag{2}$$

Now, using the change of notation from $f(x, y)$ to $F(r, \theta)$, we see that

$$f(x_k, y_k) \, \Delta A_k = F(r_k, \theta_k) r_k \, \Delta r_k \, \Delta\theta_k.$$

Hence, from (1),

$$\iint_R f(x, y) \, dA = \lim \sum_{k=1}^{N} F(r_k, \theta_k) r_k \, \Delta r_k \, \Delta\theta_k. \tag{3}$$

We now have what is necessary for expressing the double integral as an iterated integral in polar coordinates. It will be an integral of one of the forms

$$\int d\theta \int F(r, \theta) r \, dr, \qquad \int dr \int F(r, \theta) \, d\theta \tag{4}$$

with suitable limits of integration. Note that we convert $f(x, y)$ to the polar form $F(r, \theta)$ and replace $dA$ by $r \, dr \, d\theta$. The extra factor $r$ comes in by way of (2). The step from (3) to (4) is structurally just like the step from (6) to (8) in § 20-2, with the earlier roles of $x$ and $y$ now taken by $r$ and $\theta$.

The proper limits of integration are determined from a diagram. If the first integration is with respect to $r$, we select a typical $\theta$ and examine

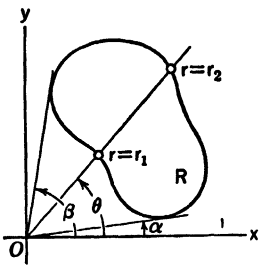



Fig. 20-12

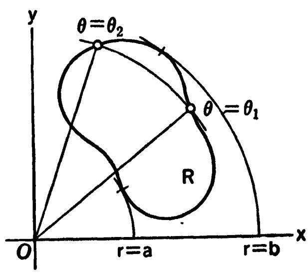



Fig. 20-13

the limits between which $r$ varies as a point crosses the region $R$ on the ray determined by $\theta$ (see Fig. 20-12). For the typical $\theta$ denote the smallest $r$ by $r_1$ and the largest $r$ by $r_2$. These values will in general be functions of $\theta$. Let the smallest and largest values of $\theta$ in $R$ be $\alpha$ and $\beta$, as shown in Fig. 20-12. Then the iterated integral is

$$\int_\alpha^\beta d\theta \int_{r_1}^{r_2} F(r, \theta) r \, dr. \tag{5}$$

The limits for the other order of integration are determined by an analogous process; the scheme is shown in Fig. 20-13. The corresponding iterated integral is

$$\int_a^b dr \int_{\theta_1}^{\theta_2} F(r, \theta) r\, d\theta. \tag{6}$$

Here $\theta_1$ and $\theta_2$ will in general be functions of $r$.

**Example 1:** A homogeneous circular lamina is bounded by the circle $r = 2a \cos \theta$ (Fig. 20-14). Find its moment of inertia about an axis perpendicular to the plane of the lamina at the origin.

We have

$$I = \iint_R (x^2 + y^2)\sigma\, dA,$$

where $\sigma$ is constant. For this problem, if we integrate first with respect to $r$, we see that for a typical $\theta$ the variation of $r$ is from 0 to $2a \cos \theta$; the total variation of $\theta$ is from $-\pi/2$ to $\pi/2$. Hence, since $\sigma$ is constant and $x^2 + y^2 = r^2$, we have

$$I = \sigma \int_{-\pi/2}^{\pi/2} d\theta \int_0^{2a \cos \theta} r^3\, dr.$$

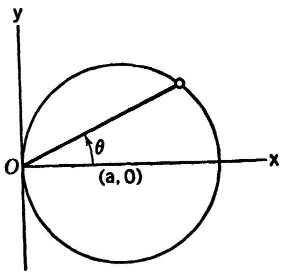



**Fig. 20-14**

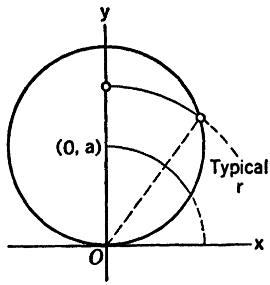



**Fig. 20-15**

The first integration yields

$$\int_0^{2a \cos \theta} r^3\, dr = \frac{1}{4}(2a \cos \theta)^4 = 4a^4 \cos^4 \theta.$$

In the second integration we can integrate from 0 to $\pi/2$ and double the result, because of symmetry. Hence, using formula 107 from the table of integrals, we have

$$I = 8a^4\sigma \int_0^{\pi/2} \cos^4 \theta\, d\theta = 8a^4\sigma \frac{1 \cdot 3}{2 \cdot 4} \frac{\pi}{2},$$

or

$$I = \frac{3\pi a^4 \sigma}{2}.$$

The mass is $M = \pi a^2 \sigma$. Hence we can write $I = \frac{3}{2} Ma^2$.

**Example 2: Consider the first quadrant portion of the region inside the circle $r = 2a \sin \theta$ and outside the circle $r = a$ (Fig. 20-15). If this region is**

regarded as a homogeneous lamina of unit density, find its first moment about the $y$-axis. This first moment is by definition the integral $\iint_R x\sigma\, dA$.

Switching to polar coordinates, we write $x = r\cos\theta$. We must also remember to put in the factor $r$ (i.e., to put $r\,dr\,d\theta$ in place of $dA$). This time we integrate first with respect to $\theta$. For a typical $r$, $\theta$ varies from $\sin^{-1}(r/2a)$ to $\pi/2$; the first of these values of $\theta$ comes from the equation $r = 2a\sin\theta$. The extreme values of $r$ are $a$ and $2a$. Hence (putting $\sigma = 1$)

$$\iint_R x\, dA = \int_a^{2a} dr \int_{\sin^{-1}(r/2a)}^{\pi/2} r^2 \cos\theta\, d\theta.$$

At the first integration we have

$$r^2 \int_{\sin^{-1}(r/2a)}^{\pi/2} \cos\theta\, d\theta = r^2\left[\sin\frac{\pi}{2} - \sin\left(\sin^{-1}\frac{r}{2a}\right)\right]$$

$$= r^2\left(1 - \frac{r}{2a}\right).$$

Then

$$\iint_R x\, dA = \int_a^{2a}\left(r^2 - \frac{r^3}{2a}\right) dr = \left(\frac{r^3}{3} - \frac{r^4}{8a}\right)\Bigg|_a^{2a}$$

$$= \left(\frac{8a^3}{3} - 2a^3\right) - \left(\frac{a^3}{3} - \frac{a^3}{8}\right) = \frac{11}{24}a^3.$$

*Moments of Inertia. The Parallel Axis Theorem*

Suppose we have a certain distribution of mass. It may consist of particles, or of matter continuously distributed in various ways (wires, laminas, solids), or of combinations of these things. Let the total mass be $M$. Let $L$ and $L_0$ be two parallel axes a distance $h$ apart, with $L_0$ passing through the center of mass of the system. Let $I$ and $I_0$ be the moments of inertia of the system about $L$ and $L_0$, respectively. Then it can be proved that

$$I = I_0 + Mh^2. \tag{7}$$

This assertion is called *the parallel axis theorem.* We shall prove it for the special case in which the total mass is spread over a lamina occupying a region $R$ in the $xy$-plane and the two axes are themselves in the $xy$-plane. The method of proof can be adapted to prove the theorem in other cases as well.

We shall choose the coordinate system so that $L$ coincides with the $y$-axis. Then

$$I = \iint_R x^2\sigma\, dA, \qquad I_0 = \iint_R (x - \bar{x})^2\sigma\, dA,$$

where $\bar{x}$ is the abscissa of the center of mass. We have to prove that

$I - I_0 = Mh^2$. Now

$$M = \iint_R \sigma\, dA, \qquad M\bar{x} = \iint_R x\sigma\, dA.$$

Then

$$I - I_0 = \iint_R [x^2 - (x - \bar{x})^2]\sigma\, dA = \iint_R (2x\bar{x} - \bar{x}^2)\sigma\, dA$$

$$= 2\bar{x}M\bar{x} - \bar{x}^2M = M\bar{x}^2.$$

Since $h = |\bar{x}|$, this result is equivalent to (7).

Here is another interesting simple fact about moments of inertia. Consider the three mutually perpendicular axes in the $xyz$-coordinate system. Suppose we have a distribution of mass in the $xy$-plane, and let its moments of inertia about the $x$-axis, the $y$-axis, and the $z$-axis, respectively, be $I_x$, $I_y$, $I_z$. Then

$$I_z = I_x + I_y. \tag{8}$$

This is easily proved, and we leave the argument to the reader. See Exercise 9. Is the result (8) valid if the mass is not all in the $xy$-plane?

*Products of Inertia*

For a lamina occupying a region $R$ in the $xy$-plane, the double integral

$$\iint_R xy\sigma\, dA \tag{9}$$

is called the *product of inertia* of the lamina relative to the coordinate axes. In certain cases this product of inertia will be zero. For example, if $\sigma$ is constant and the region $R$ is symmetric with respect to one of the coordinate axes, the integral in (9) will be zero. (Why?)

For a given lamina let us write $I_x = A$, $I_y = C$, and denote the product of inertia in (9) by $B$. If we know the values of $A$, $B$, $C$, we can compute the moment of inertia of the lamina about any axis which lies in the $xy$-plane and goes through the origin. In fact, if the equation of the axis is $y = x \tan \theta$, then the corresponding moment of inertia is

$$I = A \cos^2 \theta - 2B \sin \theta \cos \theta + C \sin^2 \theta. \tag{10}$$

This is easily proved by using the formula for distance from a point $(x, y)$ to the axis. See Exercise 10.

## EXERCISES

**1.** Find the moment of inertia about the $z$-axis of a homogeneous lamina occupying the indicated region in the $xy$-plane. Express answers in the form $I = Mk^2$, where $M$ is the mass of the lamina.

(a) $R$ defined by $x^2 + y^2 \leq a^2$.
(b) $R$ the region bounded by the cardioid $r = a(1 + \cos\theta)$.
(c) $R$ the region bounded by the lemniscate $r^2 = a^2 \cos 2\theta$.

**2.** A lamina occupies the region bounded by the circle $r = 2a \sin\theta$. The density is $\sigma = cr$, where $c$ is a constant. Locate the center of mass of the lamina.

**3.** Calculate the volume of each solid, using a double integral to express the volume, and then calculating it by an iterated integral in polar coordinates.
(a) The first octant portion of the solid sphere $x^2 + y^2 + z^2 \leq a^2$.
(b) The wedge-shaped solid inside the ellipsoid $9(x^2 + y^2) + 4z^2 = 36$, in the first octant, and between the planes $x = 0$, $x = \sqrt{3}y$.
(c) The solid cut from the sphere $x^2 + y^2 + z^2 = 4a^2$ by the cylinder $(x - a)^2 + y^2 = a^2$.
(d) The solid inside the cylinder $x^2 + y^2 = 2ay$, between the plane $z = 0$ and the cone $z = \sqrt{x^2 + y^2}$.

**4.** Locate the centroids of the plane regions described as follows.
(a) The region bounded by the cardioid $r = a(1 + \sin\theta)$.
(b) The sector of the circle $x^2 + y^2 \leq a^2$ in which $|\theta| \leq \pi/6$.
(c) The region inside the loop of $r^2 = a^2 \cos 2\theta$ on which $x \geq 0$.
(d) The region inside the first quadrant loop of $r = a \sin 2\theta$.

**5.** Let $R$ be the region in the first and second quadrants and inside both the circle $r = a$ and the cardioid $r = a(1 - \cos\theta)$.
(a) Find its mass if it is a lamina of unit density. Use the sum of two iterated integrals, integrating first with respect to $r$.
(b) Find the mass if the density is $\sigma = \sin\theta$; integrate first with respect to $\theta$.

**6.** Let $R$ be the region outside the circle $r = a$ and inside the cardioid of Exercise 5, and let its density as a lamina be $\sigma = c/r$, where $c$ is constant. Find the mass and locate the center of mass.

**7.** Find the mass of the lamina of unit density occupying the area common to the circles $r = a \sin\theta$, $r = 2a \cos\theta$.

**8.** A square lamina is bounded by the lines $x = a$, $y = a$ and the coordinate axes. If the density varies in direct proportion to the distance from the corner at the origin, find the mass. Take advantage of symmetry with respect to the line $y = x$.

**9.** Prove formula (8). As a check on (8), compute $I_x$ and $I_y$ for the lamina of Example 1 in the text (Fig. 20-14) and compare with the result found in Example 1.

**10.** (a) Prove formula (10), using the formula for distance from $(x, y)$ to the line $y \cos\theta = x \sin\theta$. (b) Using formula (10), and the meanings of $A$, $B$, $C$ in that formula, consider the ellipse $Ax^2 - 2Bxy + Cy^2 = 1$. Let $L$ be a line through the origin in the $xy$-plane, and let $R$ be the distance to the origin from where this line $L$ cuts the ellipse. Show that the moment of inertia of the given lamina about the axis $L$ is $1/R^2$. Because of this the ellipse is called the *ellipse of inertia* for the given lamina, relative to the point $O$. The result just stated shows that $I$ is smallest when $L$ coincides with the major axis of the ellipse of inertia. The axes of symmetry of the ellipse are called *principal axes of inertia* of the lamina (relative to $O$).

11. Compute the product of inertia (9) for the homogeneous lamina defined by $x^2 + y^2 \le a^2$, $y \ge 0$, $(x - a)^2 + y^2 \le a^2$.

12. Get the equation of the ellipse of inertia relative to $O$ for the lamina of density $\sigma = (x + y)^2$ occupying the region $x^2 + y^2 \le a^2$. What are the principal axes of inertia?

13. Among all axes parallel to a given line, about which one is the moment of inertia of a given mass system the least?

14. Suppose $f(x) \ge 0$ when $a \le x \le b$, $f$ being a continuous function. Consider the volume generated when the area between the curve $z = f(x)$ and the $x$-axis in the $xz$-plane, from $x = a$ to $x = b$, is revolved around the $z$-axis. Show that this volume is given by the double integral $\iint_R f(\sqrt{x^2 + y^2})\,dx\,dy$, where $R$ is the region between the circles $x^2 + y^2 = a^2$, $x^2 + y^2 = b^2$ in the $xy$-plane. Express the double integral as an iterated integral in polar coordinates and show that the result is in agreement with the shell method of finding volumes of solids of revolution, in § 11-2.

## 20-4 Mass Systems and Newton's Law

The present section is a digression from the subject of multiple integrals. We shall discuss the way in which Newton's second law of motion is applied to the study of the motion of a mass system. The discussion brings out the importance of the concept of center of mass and also the importance of the concept of moment of inertia.

We begin with the consideration of a rigid mass system which is rotating about a fixed axis. When we describe the system as rigid we mean that if we fix our attention on any two points in the mass system, the distance between these two points does not change as the whole system moves. As a consequence of the rigidity, when the system rotates, any particular point of the mass system describes a circular path about the axis of rotation, *and all points move with the same angular velocity.* A mass particle $m_k$ at distance $r_k$ from the axis moves with speed $v_k = \omega r_k$, where $\omega$ is the angular velocity. Hence its kinetic energy is

$$\tfrac{1}{2}m_k v_k^2 = \tfrac{1}{2}m_k r_k^2 \omega^2.$$

If the system consists of $n$ particles, the total kinetic energy is

$$\tfrac{1}{2}\sum_{k=1}^{n} m_k v_k^2 = \tfrac{1}{2}\omega^2 \sum_{k=1}^{n} m_k r_k^2 = \tfrac{1}{2}I\omega^2,$$

where $I$ is moment of inertia of the system relative to the axis of rotation. This expression:

$$\text{kinetic energy} = \tfrac{1}{2}I\omega^2, \tag{1}$$

is valid for the motion of all rigid mass systems rotating about a fixed axis. In the case of laminas or other types of continuously distributed mass systems, (1) is taken as a definition.

When a rigid mass system moves in such a way that there is no fixed axis of rotation, matters are more complicated. It may be, however, that each point of the mass system moves in a plane, and that the planes corresponding to different points are all parallel. The example of a sphere rolling along on a table is an illustration. In this case we may pass an axis through the center of mass of the system, at right angles to the plane in which this point moves. Then at any given instant we may speak about the angular velocity of the system relative to this axis. It is not hard to show that in this case the proper formula for kinetic energy is

$$\text{kinetic energy} = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\omega^2, \tag{2}$$

where $M$ is the total mass, $v$ is the linear speed of the center of mass, and $I$ and $\omega$ refer to the rotation of the system about the axis through the center of gravity as described. The derivation of this formula, for the case of a rigid system of particles moving in the $xy$-plane, is left to the student; see Exercise 4. We shall not consider motions of a more complicated character.

*Newton's Law and the Motion of the Center of Mass*

Consider a system consisting of $n$ particles, of masses $m_1, \cdots, m_n$. We think of a general case of motion in three dimensions, *and the system need not be rigid.* Let each mass be acted on by forces, some of which are related to the presence of the other masses. We denote the forces on $m_1$ by $\mathbf{F}_1$, a force from outside the system, and by $\mathbf{F}_{12}, \mathbf{F}_{13}, \cdots, \mathbf{F}_{1n}$. Here $\mathbf{F}_{1k}$ denotes the force on $m_1$ which is related to the presence of $m_k$. We call $\mathbf{F}_1$ an *external* force; the forces $\mathbf{F}_{1k}$ are called *internal.* On $m_2$ there will be forces $\mathbf{F}_2$ and $\mathbf{F}_{21}, \mathbf{F}_{23}, \cdots, \mathbf{F}_{2n}$, and so on for the other masses. We make the important assumption that the internal forces are equal and opposite in pairs:

$$\mathbf{F}_{12} + \mathbf{F}_{21} = 0, \qquad \mathbf{F}_{13} + \mathbf{F}_{31} = 0, \qquad \text{etc.} \tag{3}$$

This assumption is fulfilled if, for instance, the internal forces are due to gravitational attraction. Or, each pair of particles might be tied together by a cord or rod of negligible mass. Finally, we let $M$ be the total mass, and we let $\mathbf{F}$ be the vector sum of the external forces. *Then we can show, as a consequence of Newton's second law, that the center of mass of the system moves just as though it were a particle of mass $M$ acted on by the force* $\mathbf{F}$.

The proof is quite simple. If $\mathbf{R}_k$ is the position vector from the origin $O$

to the point $(x_k, y_k, z_k)$ occupied by $m_k$, then Newton's law asserts that

$$m_k \frac{d^2\mathbf{R}_k}{dt^2} = \mathbf{F}_k + \text{internal forces on } m_k.$$

If we add all the equations of this type together and take note of the relations in (3), we see that

$$m_1 \frac{d^2\mathbf{R}_1}{dt^2} + \cdots + m_n \frac{d^2\mathbf{R}_n}{dt^2} = \mathbf{F}. \tag{4}$$

Now let $(\bar{x}, \bar{y}, \bar{z})$ be the center of mass, and let $\mathbf{R}$ be the vector from $O$ to $(\bar{x}, \bar{y}, \bar{z})$. Since

$$M\bar{x} = m_1x_1 + \cdots + m_nx_n,$$

and similar formulas hold for $\bar{y}$ and $\bar{z}$, it is easy to see that

$$M \frac{d^2\mathbf{R}}{dt^2} = \mathbf{F}. \tag{5}$$

This formula can be interpreted as Newton's law for the motion of a particle of mass $M$ acted on by the force $\mathbf{F}$. Hence our earlier italicized assertion has been proved.

*The Principle of Angular Momentum*

There is another useful general theorem about finite systems of particles. For its statement we need the concept of *the moment of a vector with respect to the origin.* For this concept the vector must be thought of as based at a definite point. Let $P$ be a point, let $\mathbf{R}$ be the vector from $O$ to $P$, and let $\mathbf{A}$ be a vector based at $P$ (see Fig. 20-16). Then the moment of $\mathbf{A}$ with respect to $O$ is defined to be the cross product $\mathbf{R} \times \mathbf{A}$. For information about cross products of vectors see §18-4. If neither $\mathbf{R}$ nor $\mathbf{A}$ is $\mathbf{0}$, and if $\mathbf{R}$ and $\mathbf{A}$ are not collinear, $\mathbf{R} \times \mathbf{A}$ is a vector perpendicular to the plane of $\mathbf{R}$ and $\mathbf{A}$, and of magnitude $|\mathbf{R}||\mathbf{A}| \cos \alpha$, where $\alpha$ is the acute angle which $\mathbf{A}$ makes with a line perpendicular to $\mathbf{R}$ in the plane of $\mathbf{R}$ and $\mathbf{A}$ (see Fig. 20-16). In other words, the magnitude of $\mathbf{R} \times \mathbf{A}$ is the product of the length of $\mathbf{R}$ and the length of the component of $\mathbf{A}$ at right angles to $\mathbf{R}$ in the plane of $\mathbf{R}$ and $\mathbf{A}$. This explains why we call $\mathbf{R} \times \mathbf{A}$ the *moment* of $\mathbf{A}$, if we think of $\mathbf{A}$ as a force and $\mathbf{R}$ as a lever arm.

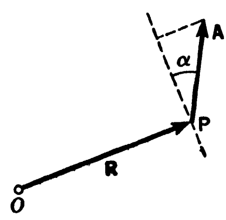


Fig. 20-16

Now consider the system of mass particles $m_1, \cdots, m_k$, just described in the preceding discussion of Newton's law. The *linear momentum* of the mass $m_k$ is defined to be the vector

$$m_k \frac{d\mathbf{R}_k}{dt}$$

(i.e., mass times vector velocity). If we think of this linear momentum as a vector based at the point occupied by $m_k$, the moment of this momentum vector is called the *angular momentum* of $m_k$. Thus the angular momentum of $m_k$ is

$$\mathbf{R}_k \times m_k \frac{d\mathbf{R}_k}{dt}.$$

The total angular momentum of the system of particles is then defined to be the sum

$$\mathbf{H} = \sum_{k=1}^{n} \mathbf{R}_k \times m_k \frac{d\mathbf{R}_k}{dt}. \tag{6}$$

Let us compute the derivative of **H** with respect to time. If we have to differentiate a cross product $\mathbf{A} \times \mathbf{B}$, the appropriate formula is

$$\frac{d}{dt}(\mathbf{A} \times \mathbf{B}) = \mathbf{A} \times \frac{d\mathbf{B}}{dt} + \frac{d\mathbf{A}}{dt} \times \mathbf{B}.$$

This is easily worked out. Then

$$\frac{d}{dt}\left[\mathbf{R}_k \times m_k \frac{d\mathbf{R}_k}{dt}\right] = \mathbf{R}_k \times m_k \frac{d^2\mathbf{R}_k}{dt^2},$$

because

$$\frac{d\mathbf{R}_k}{dt} \times \frac{d\mathbf{R}_k}{dt} = 0$$

as a result of the fact that the cross product of any vector with itself is zero. Therefore we find that

$$\frac{d\mathbf{H}}{dt} = \sum_{k=1}^{n} \mathbf{R}_k \times m_k \frac{d^2\mathbf{R}_k}{dt^2}. \tag{7}$$

If we combine (7) with Newton's law for each particle, we obtain an important relation between the rate of change of **H** and the forces acting on the system. In order to get this relation in a simple form we make a further assumption about the internal forces. In addition to the assumption [see (3)] that they are equal and opposite in pairs, we assume that when $\mathbf{F}_{ij}$ is based at $m_i$, it lies along the line joining $m_i$ and $m_j$. In other words, we assume that the force exerted on $m_i$ by $m_j$ is either a pull toward $m_j$ or a push away from it. The effect of our assumptions on the internal forces is then that we have relations of the type

$$\mathbf{R}_1 \times \mathbf{F}_{12} + \mathbf{R}_2 \times \mathbf{F}_{21} = (\mathbf{R}_1 - \mathbf{R}_2) \times \mathbf{F}_{12} = 0, \tag{8}$$

because $\mathbf{F}_{21} = -\mathbf{F}_{12}$ and the vectors $\mathbf{R}_1 - \mathbf{R}_2$ and $\mathbf{F}_{12}$ are collinear. See Fig. 20-17. As a result, when we combine Newton's law with (7), the effect of the internal forces is canceled out, and we obtain

$$\frac{d\mathbf{H}}{dt} = \sum_{k=1}^{n} \mathbf{R}_k \times \mathbf{F}_k. \tag{9}$$

This equation is called *the principle of angular momentum* for the system. Its great usefulness is in studying the turning or rotation of the system. We mention without proof that this principle also applies if the point $O$ is taken to be the center of mass of the system, instead of being a point fixed in space. In that case $\mathbf{H}$ is called the angular momentum relative to the center of mass.

To see the meaning of the principle of angular momentum in an important special case, suppose that the masses are all in the $xy$-plane, that

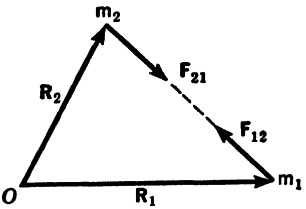



**Fig. 20-17**

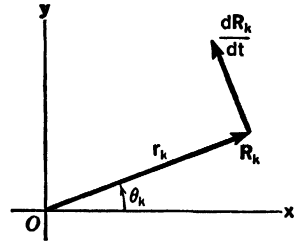



**Fig. 20-18**

the system is rigid, and that it is rotating about the $z$-axis. Then it is easy to compute the angular momentum of the system. In this case each position vector $\mathbf{R}_k$ is of constant length $r_k$; the velocity vector $d\mathbf{R}_k/dt$ is at right angles to $\mathbf{R}_k$ and has length $r_k\, d\theta_k/dt$ (see Fig. 20-18). Hence the angular momentum of $m_k$ is a vector parallel to the $z$-axis, given by

$$\mathbf{R}_k \times m_k \frac{d\mathbf{R}_k}{dt} = \left(m_k r_k^2 \frac{d\theta_k}{dt}\right)\mathbf{k}.$$

The rigidity of the system implies that $d\theta_k/dt$ is the same for all values of $k$, and hence $d\theta_k/dt = d\theta/dt$, where $\theta$ is the angular coordinate of a ray through $O$ and any selected point of the moving rigid system. Thus we see that

$$\mathbf{H} = \left(\sum_{k=1}^{n} m_k r_k^2 \frac{d\theta}{dt}\right)\mathbf{k} = \left(I \frac{d\theta}{dt}\right)\mathbf{k},$$

where $I$ is the moment of inertia of the system about the $z$-axis. If we now take components on both sides of (9), we get the principle of angular momentum in the form

$$I \frac{d^2\theta}{dt^2} = \text{algebraic sum of the force-moments about the axis of rotation} \tag{10}$$

for the planar rigid system.

**Example 1:** Suppose that the $xy$-plane is vertical and that external forces are just those due to gravity, so that $\mathbf{F}_k$ is a force of magnitude $m_k g$ in the downward vertical direction.

If we take the $x$-axis positively downward (see Fig. 20-19), the algebraic sum of the force moments is

$$-\sum_{k=1}^{n} m_k g y_k = -Mg\bar{y},$$

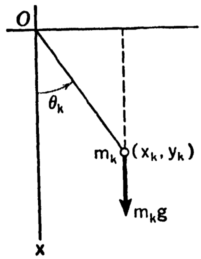



Fig. 20-19

where $M$ is the total mass, and the center of mass is at $(\bar{x}, \bar{y})$. Hence for this case

$$I\frac{d^2\theta}{dt^2} = -Mg\bar{y}. \tag{11}$$

By considering a lamina as the limiting case of a finite system of particle, we regard (11) as applicable to the case of a lamina hung up by a horizontal axis and free to oscillate in its own plane. We call the oscillating system a *compound pendulum.* If the system is a single mass $m$ a distance $r$ from the origin, (11) becomes

$$mr^2\frac{d^2\theta}{dt^2} = -mgr\sin\theta,$$

or

$$\frac{d^2\theta}{dt^2} = -\frac{g}{r}\sin\theta. \tag{12}$$

This system is called a *simple* pendulum of length $r$.

For the compound pendulum, let us write $I = Mk^2$, where $k$ is called the radius of gyration. Let $(l, \theta)$ be polar coordinates of the center of mass, so that $\bar{y} = l\sin\theta$. Then (11) can be written in the form

$$\frac{d^2\theta}{dt^2} = -\frac{gl}{k^2}\sin\theta. \tag{13}$$

A comparison of (12) and (13) shows that the compound pendulum will oscillate like a simple pendulum of length $r = k^2/l$.

**Example 2:** Suppose a circular lamina (a coin, for example) rolls down an incline without slipping. Suppose the mass distribution is such that the center of mass of the lamina is at the center of the circle. Study the way in which the motion is influenced by the moment of inertia of the lamina about a horizontal axis through its center.

We place the axes as shown in Fig. 20-20; the incline makes an angle $\alpha$

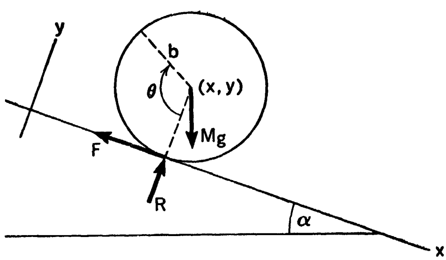



Fig. 20-20

with the horizontal. The radius of the lamina is $b$, and the center is at $(x, y)$. The external forces are: gravity $Mg$, a frictional force $F$, and a normal reaction $R$, as shown in Fig. 20-20. The no-slipping condition is expressed by the requirement $x = b\theta +$ a constant.

For the motion of the center of mass we have, applying (5),

$$M\frac{d^2x}{dt^2} = Mg\sin\alpha - F. \tag{14}$$

Next we apply the principle of angular momentum in the form (10), with $I$ and the moments calculated relative to the horizontal axis through the center of mass. The result is

$$I\frac{d^2\theta}{dt^2} = bF; \tag{15}$$

note that $F$ is the only force which produces a nonzero moment. We can now use (14) and (15) to eliminate $F$. Since $d^2x/dt^2 = b\,d^2\theta/dt^2$, we obtain

$$Mb\frac{d^2\theta}{dt^2} = Mg\sin\alpha - \frac{I}{b}\frac{d^2\theta}{dt^2},$$

or, putting $I = Mk^2$,

$$\frac{d^2\theta}{dt^2} = \frac{bg\sin\alpha}{b^2 + k^2}. \tag{16}$$

Thus the disk rolls with constant angular acceleration. The influence of the moment of inertia is felt through the term $k^2$ in the denominator.

## EXERCISES

1. A homogeneous circular lamina of radius $a$ is hung up on a horizontal axis $b$ units from the center $(0 < b \leq a)$, thus making a compound pendulum.
   (a) Find the length of the simple pendulum which will oscillate in the same manner.
   (b) For what value of $b$ is this "equivalent" simple pendulum shortest?

2. A square lamina of diagonal $c$ and an equilateral triangular lamina of altitude $h$ each oscillate about a horizontal axis perpendicular to their planes at a vertex. Find $c$ and $h$ if these compound pendulums are to oscillate in exactly the same manner as the circular lamina of Exercise 1, with $b = a = 1$ (diameter 2).

3. A uniform circular hoop of radius 1 (all the mass in the circumference) and a uniform circular lamina of radius $r$ roll down the same incline, both without slipping. What is the value of $r$ if they experience the same angular acceleration?

4. Prove the kinetic energy formula (2) for the case of a rigid system of particles moving in the $xy$-plane. *Suggestion:* Let a rectangular system of coordinates $(u, v)$ be established with origin at the center of mass, the $u$-axis being parallel to the $x$-axis, and the $v$-axis parallel to the $y$-axis. If the mass $m_k$ has coordinates $(x_k, y_k)$ and $(u_k, v_k)$ in the two systems,

explain why

$$\sum_{k=1}^{n} m_k \frac{du_k}{dt} = 0,$$

with a similar relation for the $v_k$'s. Then show that the kinetic energy is

$$\frac{1}{2}M\left[\left(\frac{d\bar{x}}{dt}\right)^2 + \left(\frac{d\bar{y}}{dt}\right)^2\right] + \frac{1}{2}\sum_{k=1}^{n} m_k\left[\left(\frac{du_k}{dt}\right)^2 + \left(\frac{dv_k}{dt}\right)^2\right],$$

and finally explain how this yields (2).

## 20-5 Surface Integrals

Let $f$ be a continuous function of $x$ and $y$, defined when $(x, y)$ is a point of a specified region $R$ in the $xy$-plane. Consider the locus of points $(x, y, z)$, where $z = f(x, y)$ and $(x, y)$ is in $R$. This locus, let us call it $L$, is a surface. We wish to talk about the area of a surface represented in this way. In order to be able to define this area and compute it in a satisfactory way, we shall assume that the function $f$ has continuous first partial derivatives.

We now refer back to Fig. 20-3, near the beginning of this chapter. The rectangular column built on the rectangular cell of area $\Delta A_i$ in $R$ cuts out a piece from the surface $L$, which appears as the top of the solid in Fig. 20-3. In order to define the area of $L$ we propose to work out an expression which seems satisfactory as an approximate representation of the area of this piece of $L$. Then the area of $L$ will be defined as the limit of the sum of all the approximations, corresponding to all the rectangular cells in the base-region $R$. In this way the area $S$ of $L$ will be obtained as a certain double integral.

The method of getting what seems intuitively to be a reasonable approximation of the area of the piece of surface in question is this: we select a point on the piece of surface, draw the tangent plane there, and compute the area of the piece of this tangent plane which is directly above the base area $\Delta A$ (see Fig. 20-21). If $\gamma$ is the acute angle which the normal to this tangent plane makes with the $z$-axis, the area $\Delta S$ of the piece of tangent plane is related to $\Delta A$ by the formula

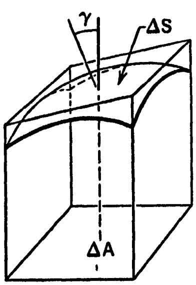


Fig. 20-21

$$\Delta S \cos\gamma = \Delta A, \quad \text{or} \quad \Delta S = \sec\gamma\,\Delta A.$$

Hence the total area of the surface is defined as the limit of the sum of all the expressions $\sec\gamma\,\Delta A$ coming from the various cells into which $R$ is divided. Thus, the *definition* of the surface area culminates in the formula

$$S = \iint_R \sec\gamma\,dA. \tag{1}$$

In order to use the formula, we must express sec $\gamma$ in terms of $x$ and $y$. Now the normal to the surface has the direction

$$\frac{\partial f}{\partial x} : \frac{\partial f}{\partial y} : -1,$$

and hence, since $\gamma$ is acute,

$$\sec \gamma = \left[1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right]^{1/2}. \tag{2}$$

If the surface $L$ is represented for us by an equation $F(x, y, z) = 0$, where $F$ has continuous derivatives and $\partial F/\partial z \neq 0$, we have a different formula for sec $\gamma$. In this case the direction of the normal is

$$\frac{\partial F}{\partial x} : \frac{\partial F}{\partial y} : \frac{\partial F}{\partial z},$$

and so

$$\sec \gamma = \frac{\left[\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F^2}{\partial z}\right)\right]^{1/2}}{\left|\frac{\partial F}{\partial z}\right|}. \tag{3}$$

**Example 1:** Let $R$ be the region which is the first quadrant portion of the part of the interior of the ellipse $x^2 + 4y^2 = 4a^2$ cut off between the lines $x = 0$, $x = a$. Let the equation of $L$ be $z = \sqrt{4a^2 - x^2 - y^2}$, so that $L$ is part of the sphere $x^2 + y^2 + z^2 = 4a^2$. See Fig. 20-22.

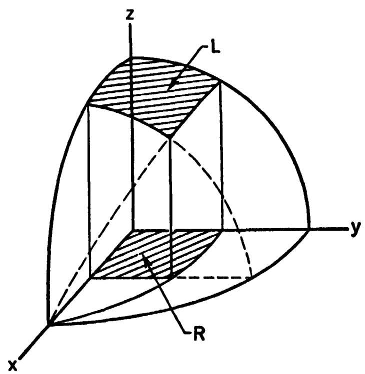



**Fig. 20-22**

If we compute sec $\gamma$ by (2), we have (some details are omitted)

$$\frac{\partial f}{\partial x} = \frac{-x}{\sqrt{4a^2 - x^2 - y^2}}, \qquad \sec \gamma = \frac{2a}{\sqrt{4a^2 - x^2 - y^2}}.$$

Alternatively, we could use (3), starting from

$$F(x, y, z) = x^2 + y^2 + z^2 - 4a^2, \quad \frac{\partial F}{\partial x} = 2x, \quad \text{etc.},$$

whence $$\sec\gamma = \frac{\sqrt{4x^2 + 4y^2 + 4z^2}}{2|z|} = \frac{2a}{\sqrt{4a^2 - x^2 - y^2}}.$$

Thus the required area is

$$S = \iint_R \frac{2a}{\sqrt{4a^2 - x^2 - y^2}}\, dA = 2a \int_0^a dx \int_0^{y_1} \frac{dy}{\sqrt{4a^2 - x^2 - y^2}},$$

where $y_1 = \sqrt{4a^2 - x^2}/2$. Now

$$\int_0^{y_1} \frac{dy}{\sqrt{4a^2 - x^2 - y^2}} = \sin^{-1} \frac{y}{\sqrt{4a^2 - x^2}} \Bigg|_{y=0}^{y=y_1} = \sin^{-1}\left(\frac{1}{2}\right) = \frac{\pi}{6}.$$

Then $$S = 2a \int_0^a \frac{\pi}{6}\, dx = 2a^2 \frac{\pi}{6} = \frac{\pi a^2}{3}.$$

In some cases it is possible to compute surface areas in a different way, by using in a suitable manner the arc lengths of certain curves on the surface. This technique is well illustrated by the following example, in which we use angles for longitude and colatitude on a sphere.

**Example 2:** Consider the sphere $x^2 + y^2 + z^2 = a^2$.

Let $\theta$, $\phi$ be angles as indicated in Fig. 20-23. Observe that $OP_0 = a$,

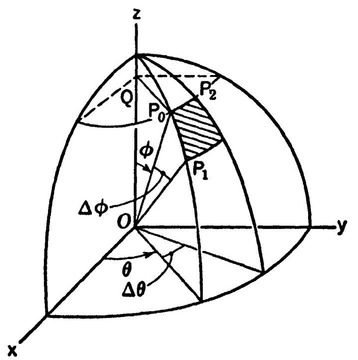



**Fig. 20-23**

$QP_0 = a \sin\phi$. Hence the length of the arc $P_0P_1$ is $a\Delta\phi$ and the arc $P_0P_2$ is $a \sin\phi\, \Delta\theta$. This suggests that the area of the shaded patch on the sphere is approximately $a^2 \sin\phi\, \Delta\theta\, \Delta\phi$, and that the area of any portion of the sphere can be found by evaluating

$$\iint a^2 \sin\phi\, d\theta\, d\phi \tag{4}$$

as an iterated integral with appropriate limits. For the entire sphere the limits would be as shown:

$$a^2 \int_0^{2\pi} d\theta \int_0^{\pi} \sin\phi\, d\phi.$$

Evaluation of this yields $4\pi a^2$, the correct result for the area of the sphere.

It can be shown that the use of (4) is consistent with the use of (1) for areas on the sphere.

In some of the exercises it is indicated how one may use integrals to compute moments of inertia and centers of mass for surfaces thought of as curved laminas.

## EXERCISES

**1.** Find the area of the portion of each surface as described.
(a) The part of $bz = x^2 - y^2$ inside the cylinder $x^2 + y^2 = a^2$ ($a$ and $b > 0$).
(b) The part of the cylinder $y^2 + z^2 = 4a^2$ on which $z \geq 0$ and $0 \leq y \leq a - x$, $0 \leq x \leq a$.
(c) The part of the cylinder $z^2 = 8x$ inside the prismatic column bounded by the planes $y = 0$, $x = 1$, and the cylinder $x^2 = 4y$.
(d) The area of the part of the cone $z^2 = x^2 + y^2$ inside the cylinder $x^2 + y^2 = 2ay$.
(e) The area cut from the sphere $x^2 + y^2 + z^2 = a^2$ by the cylinder $x^2 + y^2 = ax$.

**2.** Follow the direction of Exercise 1.
(a) The area of the part of the cylinder $y^2 + z^2 = a^2$ inside the cylinder $x^2 = a(y + a)$.
(b) The part of the cylinder $y^2 + z^2 = a^2$ on which $z \geq 0$ and $0 \leq x \leq y \leq a$.
(c) The area of the portion of the plane $M$ which is in the first octant and inside the elliptic cylinder $b^2x^2 + a^2y^2 = a^2b^2$, if the plane $M$ is determined by the points $(0, 0, 0)$, $(a, 0, b)$, $(0, b, a)$.
(d) The area cut from the sphere $x^2 + y^2 + z^2 = 4az$ by a cylinder with its elements parallel to the $z$-axis and passing through the first-quadrant loop of the curve $r = 2a \sin 2\theta$.
(e) The total area of the part of the cylinder $x^2 + z^2 = a^2$ inside the cylinder $x^2 + y^2 = ax$.

**3.** Consider a plane $M$ with direction of normal $\cos\alpha : \cos\beta : \cos\gamma$, where $0 < \gamma < \pi/2$. Consider a region in this plane, of area $S$, and let $A$ be the area of the region in the $xy$-plane obtained by projection parallel to the $z$-axis. Explain why $S \cos\gamma = A$ by an argument of the following type: Make a partition in the plane $M$ by two sets of lines, one set parallel to the $xy$-plane, and one set at right angles to the first set. Explain by simple geometry why $\Delta S \cos\gamma = \Delta A$ if $\Delta S$ is the area of one of the rectangular cells in this partition and $\Delta A$ is the area of the projection of the cell. How does the general result $S \cos\gamma = A$ follow?

**4.** (a) A right circular cone has semivertical angle $\phi$. An area $S$ on the cone is projected orthogonally onto a plane perpendicular to the axis of the cone. Show that the area of the projection is $S \sin\phi$. (b) Find the total area of the part of the cone $3(x^2 + z^2) = y^2$ inside the cylinder $x^2 + z^2 = 2az$.

5. Find the area of the part of the paraboloid $4z = x^2 + y^2$ which is inside the cylinder with elements parallel to the $z$-axis and passing through the leminiscate $r^2 = 4\cos 2\theta$ in the $xy$-plane.

6. Find the area cut from the cone $x^2 + y^2 = z^2$ by a long triangular prism whose faces are the planes $z = 2$, $x = 0$, $z = x + 1$. Assume the prism extends from $y = -3$ to $y = 3$.

7. Find the area of the portion of the sphere $x^2 + y^2 + z^2 = a^2$ inside the triangular prism whose faces are the planes $y = 0$, $x = y$, $x = a/\sqrt{2}$. Assume the prism extends from $z = -2a$ to $z = 2a$.

8. (a) If the curve $z = f(x)$ in the $xy$-plane is revolved around the $x$-axis, what is the equation of the resulting surface of revolution? Assume $f$ continuous, with continuous derivative, and $f(x) \geq 0$.
(b) Use the double integral (1) to find the area of the part of the surface of revolution in the first octant and between the planes $x = a$, $x = b$, and show in this way that the complete area of the surface of revolution between these planes is $2\pi \int_a^b f(x)\sqrt{1 + [f'(x)]^2}\,dx$. This shows that our present method is consistent with the method of § 11-4 for surfaces of revolution.

9. (a) If the surface of revolution described in Exercise 8 is regarded as a curved lamina whose density $\sigma$ is a function of $x$ only, show that its moment of inertia about the $x$-axis is given by

$$I = 2\pi \int_a^b [f(x)]^3\sqrt{1 + [f'(x)]^2}\,\sigma\,dx.$$

(b) Find the moment of inertia of a homogeneous lamina in the form of a spherical surface, about a diameter. Let the radius be $r$.
(c) Find the moment of inertia of a homogeneous lamina in the form of a right circular cone (lateral surface only) of altitude $h$ and radius of base $r$.

10. A curve is defined on the first octant portion of the sphere $x^2 + y^2 + z^2 = a^2$ by the equation $6\phi - 2\theta = \pi$, where $\theta$ and $\phi$ are the angles in Fig. 20-23. Find the area of the part of the first-octant surface of the sphere between this curve and the equatorial plane $z = 0$.

11. The first-octant surface of the sphere $x^2 + y^2 + z^2 = a^2$ is a lamina of density $\sigma = x$. Locate its center of mass. *Hint:* The mass is the double integral $\iint \sigma a^2 \sin\phi\,d\theta\,d\phi$ with limits appropriate to the part of the sphere being considered.

## 20-6 Triple Integrals

The definition of a triple integral follows the same kind of pattern that was laid out in the definition of a double integral in § 20-1. On that account we shall be rather brief in defining a triple integral.

We suppose that $T$ is a region of the three-dimensional space and that $f$ is a continuous function of $(x, y, z)$ defined in $T$. We wish to define

$$\iiint_T f(x, y, z)\, dx\, dy\, dz. \tag{1}$$

We suppose that $T$ is a comparatively simple type of region, with a boundary formed by surfaces of a sort which will be typically illustrated in the examples and exercises of following sections. A typical general case would be that in which $T$ consists of all $(x, y, z)$ such that $(x, y)$ is in a plane region $R$ of the type considered in connection with double integrals and $h_1(x, y) \leq z \leq h_2(x, y)$, where $h_1$ and $h_2$ are continuous functions such that $h_1(x, y) < h_2(x, y)$ at points in the interior of $R$.

We divide the space in $T$ into rectangular blocks (we call them cells) in a manner which is the three-dimensional counterpart of the sort of scheme shown in Fig. 20-2. Let $\Delta V_i$ be the volume of the $i$th cell; let $(x_i, y_i, z_i)$ be a point in the cell. Consider the sum of all the products $f(x_i, y_i, z_i)\,\Delta V_i$, and the limit of this sum as the maximum cell dimension approaches zero. The limit of this sum is defined to be the triple integral (1). Sometimes we use the alternative notation

$$\iiint_T f(x, y, z)\, dV.$$

Triple integrals may be used to formulate physical quantities just as was done with double integrals. If we conceive of the region $T$ as being filled with matter of density $\sigma(x, y, z)$, then the total mass is

$$M = \iiint_T \sigma(x, y, z)\, dV.$$

Likewise we can express as integrals the first moments of this mass with respect to each coordinate plane, and thus locate the center of mass $(\bar{x}, \bar{y}, \bar{z})$. The formulas are analogous to (8) in § 20-1. Moments of inertia also may be formulated as triple integrals.

*Gravitational Attraction*

Newton's law of gravitation, in its simplest form, states that one mass particle exerts on another mass particle a force of attraction whose magnitude is directly proportional to the two masses involved and inversely proportional to the square of the distance between them. That is, if mass $M$ is at $P$ and mass $m$ is at $Q$, then the force of attraction on $M$ by $m$, if based at $P$, appears as a vector having the direction from $P$ to $Q$, of length

$$\lambda \frac{mM}{(PQ)^2},$$

where $PQ$ is the distance from $P$ to $Q$ and $\lambda$ is a constant depending only on the units of mass, force, and distance. In our work we shall not be concerned with the value of $\lambda$ for particular systems of units.

The generalized form of Newton's law, for the attraction which a distribution of matter exerts on a mass particle, is expressed in terms of integrals. Imagine a mass particle $m$ at $Q$, and consider the gravitational force exerted on it by a mass of density $\sigma(x, y, z)$ distributed throughout the region $T$. We envisage the situation in which the mass distribution is replaced by a large number of particles, by dividing $T$ into cells and concentrating the mass of each cell at a point within it. See the schematic

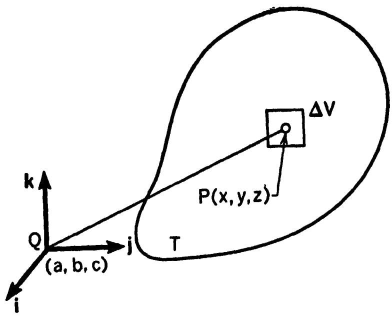



**Fig. 20-24**

diagram in Fig. 20-24. The mass at $P$ is $\sigma(x, y, z)\,\Delta V$. Hence the force in question is the vector

$$\frac{\lambda m\sigma(x, y, z)\,\Delta V}{(PQ)^2}\,\mathbf{u}, \tag{2}$$

where $\mathbf{u}$ is a unit vector with the direction of the line from $Q$ to $P$. That is,

$$\mathbf{u} = \frac{x-a}{r}\mathbf{i} + \frac{y-b}{r}\mathbf{j} + \frac{z-c}{r}\mathbf{k},$$

where $P$ is $(x, y, z)$, $Q$ is $(a, b, c)$, and $PQ = r$. This discussion makes it reasonable to formulate the statement of Newton's law in integral form: The force on $m$ at $Q$ due to all the mass in $T$ is

$$\mathbf{F} = \lambda m \iiint_T \frac{\sigma(x, y, z)}{(PQ)^3}\,[(x-a)\mathbf{i} + (y-b)\mathbf{j} + (z-c)\mathbf{k}]\,dV. \tag{3}$$

The meaning of this vector triple integral is obvious. To compute $\mathbf{F}$ we compute each component as a scalar triple integral. The $x$-component, for example, is

$$\lambda m \iiint_T \frac{x-a}{(PQ)^3}\,\sigma\,dV.$$

## 20-7 Threefold Iterated Integrals

To evaluate a triple integral by integrations with respect to $x$, $y$, and $z$ we need an argument which corresponds, for three dimensions, to the discussion of double integrals in connection with formula (6) and Fig. 20-8 in § 20-2. We shall omit any attempt to go into all the details. The result is that, if the boundary of $T$ is of a suitable nature, the triple integral

$$\iiint_T f(x, y, z)\, dx\, dy\, dz$$

can be expressed as an iterated integral of the form

$$\int_a^b dx \int_{g_1(x)}^{g_2(x)} dy \int_{h_1(x,y)}^{h_2(x,y)} f(x, y, z)\, dz \tag{1}$$

with suitable limits of integration. The way in which these limits of integration are found is illustrated in Fig. 20-25. In this diagram $T$ is a pyram-

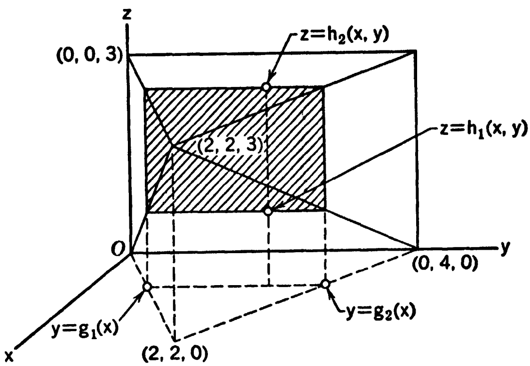



**Fig. 20-25**

idal region bounded by the planes $x = 0$, $y = x$, $x + y = 4$ (all parallel to the $z$-axis) and by the planes $2z = 3x$, $z = 3$. The base of the pyramid is a rectangle in the plane $x = 0$, and the apex is at $(2, 2, 3)$. In this case the integral (1) becomes

$$\int_0^2 dx \int_x^{4-x} dy \int_{3x/2}^{3} f(x, y, z)\, dz. \tag{2}$$

With this order of integration one begins by drawing a typical line parallel to the $z$-axis and noting the upper and lower extremities of the segment of this line within $T$. In this way we find the $z$-limits, which are functions of $x$ and $y$. Then we can project the whole region $T$ down onto the $xy$-plane and call the resulting plane region $R$. The two remaining integrations have limits just as in the case of a double integral over $R$. Alternatively, we can think of the plane section through $T$ made by the typical plane $x$ = constant. The $z$-limits and $y$-limits are then those of a double integral with respect to $y$ and $z$ over this plane section. The last integration, with respect to $x$, involves taking all possible values of $x$ in $T$, from the algebraically smallest to the algebraically largest. The limits in (2) can be obtained by either of these methods from Fig. 20-25.

There are six possible orders for an iterated integral with respect to $x$, $y$, $z$. In some cases a problem may be much simpler with one of these orders than with some of the others.

**Example:** Locate the centroid of the tetrahedron cut from the first octant by the plane with intercepts $a$, $b$, $c$ (all positive) on the axes of $x$, $y$, $z$, respectively.

The volume of the tetrahedron is $V = abc/6$. To find $\bar{z}$ we have

$$V\bar{z} = \iiint_T z\,dV.$$

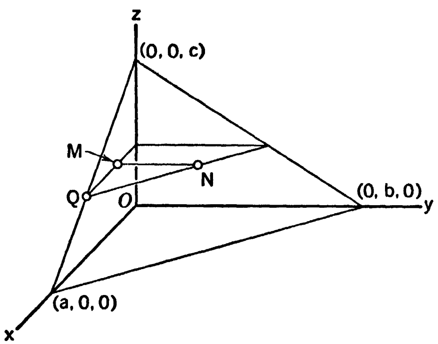



**Fig. 20-26**

To find the limits of integration we refer to Fig. 20-26 and use the equation of the plane. We integrate with respect to $y$, $x$, $z$, in that order. The equation of the slanting plane is

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1.$$

The line $MN$ is one with $x$ and $z$ constant. In terms of these constant values of $x$ and $z$, $y = b(1 - x/a - z/c)$ at $N$. To get coordinates of $Q$ we put $y = 0$ but keep the same value of $z$ as on the line $MN$. Thus $x = a(1 - z/c)$ at $Q$. We then have

$$\begin{aligned}\frac{abc}{6}\bar{z} &= \int_0^c dz \int_0^{a(1-z/c)} dx \int_0^{b(1-x/a-z/c)} z\,dy \\ &= \int_0^c bz\,dz \int_0^{a(1-z/c)} \left(1 - \frac{x}{a} - \frac{z}{c}\right) dx.\end{aligned}$$

The result of the $x$-integration is

$$\frac{-a}{2}\left(1-\frac{x}{a}-\frac{z}{c}\right)^2\Big|_{x=0}^{x=a(1-z/c)} = \frac{a}{2}\left(1-\frac{z}{c}\right)^2.$$

Thus
$$\frac{abc}{6}\bar{z} = \frac{ab}{2}\int_0^c z\left(1-\frac{z}{c}\right)^2 dz = \frac{abc^2}{24}.$$

The details of the last integral are left to the student. The final result is $\bar{z} = c/4$. By the symmetry of the situation we conclude that the centroid is at $(a/4, b/4, c/4)$.

## EXERCISES

**1.** The first-octant portion of the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} = 1$ is filled with matter of constant density. Find the mass and locate the center of mass.

**2.** The first octant portion of the solid inside the cylinder $y^2 + z^2 = a^2$ and between the planes $x = 0$, $x = b$ has density $\sigma = y$. Find the mass and locate the center of mass. Set up the limits of integration for all six orders of integration. Then choose a convenient order for calculating each of the needed triple integrals.

**3.** Find the moment of inertia, about the $z$-axis, of the homogeneous tetrahedron bounded by the planes $x = 0$, $y = 0$, $z = 1$, $x + y = z$. Set up the limits of integration in this case for all six orders of integration.

**4.** The cube bounded by the planes $x = 0$, $x = 1$, $y = 0$, $y = 1$, $z = 0$, $z = 1$ is filled with matter of density $\sigma = xz$. Find the $y$-component of the attraction which the cube exerts on a unit mass at the origin.

**5.** Let $T$ be the tetrahedron with vertices at $(0, 0, 0)$, $(5, 0, 0)$, $(0, 3, 0)$, $(0, 0, 4)$.
(a) Calculate the triple integral of $x^2$ over $T$, integrating in the order $z, y, x$.
(b) Calculate the triple integral of $y$ over $T$, integrating in the order $x, z, y$.
(c) Calculate the triple integral of $z^2$ over $T$, integrating in the order $y, x, z$.

**6.** Let $T$ be the tetrahedron with vertices $(0, 0, 0)$, $(1, 1, 0)$, $(0, 1, 0)$, $(0, 1, 1)$.
(a) Calculate the triple integral of $z$ over $T$, integrating in the order $x, y, z$.
(b) Calculate the triple integral of $yz$ over $T$, selecting a convenient order of integration.
(c) Set up the limits of integration for all five orders of integration besides the order in (a).

**7.** A homogeneous solid is bounded by the plane $z = 0$ and the paraboloid $b^2cx^2 + a^2cy^2 + a^2b^2z = a^2b^2c$. (a) Find its mass and center of mass. Set up the limits for all six orders of integration and then select a convenient order for the evaluation of the various triple integrals you need. (b) Locate the center of mass of the first-octant portion of the solid.

8. Find the moment of inertia about the $x$-axis of the tetrahedron of Fig. 20-26, if it is of constant density. Integrate first with respect to $x$.

9. Find the following masses, the solid being as described, and the density as specified.
(a) $x^2 + y^2 + z^2 \leq a^2$, $x \geq 0$, $y \geq 0$, $z \geq 0$, $\sigma = xy$.
(b) $x^2 + y^2 \leq a^2$, $0 \leq z \leq h$, $y \geq 0$, $\sigma = yz$.
(c) $4x^2 + 9y^2 \leq 36z^2$, $0 \leq z \leq 1$, $x \geq 0$, $y \geq 0$, $\sigma = xz$.

10. A solid lying in the first octant is bounded below by the plane $by = az$, above by the plane $z = b$, and laterally by the planes $x = 0$, $y = 0$ and the cylinder $x^2 + y^2 = a^2$.
(a) Set up the limits of integration for the two iterated integrals over this solid in which the first integration is with respect to $x$, and likewise for the two in which the first integration is with respect to $z$. What is different about the case when the first integration is with respect to $y$?
(b) Calculate the integral of $x$ over the solid, choosing a convenient order of integration.
(c) Calculate the integral of $y$ over the solid.

## 20-8 Cylindrical Coordinates

Cylindrical coordinates in space are a combination of plane polar coordinates and a linear coordinate along an axis perpendicular to the plane. It is customary to use polar coordinates $(r, \theta)$ in the $xy$-plane along with the usual $z$-coordinate (see Fig. 20-27). It is sometimes advantageous to calculate a triple integral as an iterated integral in cylindrical coordinates. For this purpose we replace $dV$ (or $dx\,dy\,dz$) by $r\,dr\,d\theta\,dz$, convert the integrand $f(x, y, z)$ to an expression $F(r, \theta, z)$, and integrate. The reason for $r\,dr\,d\theta\,dz$ in place of $dV$ is the same as the reason for $r\,dr\,d\theta$ in place of

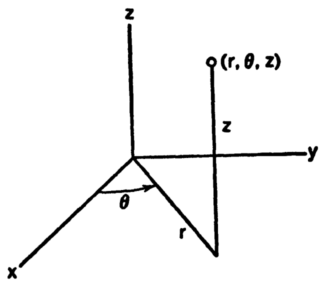



Fig. 20-27

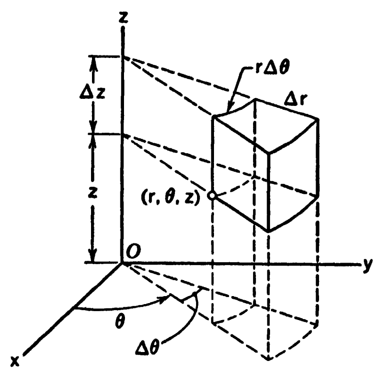



Fig. 20-28

$dA$ in double integrals (see § 20-3). If we think of dividing space into cells in a natural way, using cylindrical coordinates, we obtain a typical cell, or "volume element" as shown in Fig. 20-28. The lengths of the three mutually perpendicular edges of the cell issuing from the point $(r, \theta, z)$ are $\Delta r$, $r\,\Delta\theta$, $\Delta z$, so the volume of the cell is approximately $r\,\Delta r\,\Delta\theta\,\Delta z$.

The finding of limits of integration for iterated integrals in cylindrical coordinates is the same in principle as for iterated integrals in rectangular coordinates.

**Example:** Find the moment of inertia of a homogeneous solid sphere about a diameter.

We take the sphere to be $x^2 + y^2 + z^2 \leq a^2$, and compute the moment of inertia about the $z$-axis. Clearly it suffices to deal with the hemisphere in which $z \geq 0$ and double the result. Hence, with $T$ this hemispherical region,

$$I = 2 \iiint\limits_T (x^2 + y^2)\sigma\, dV.$$

Here $\sigma$ is constant and $M = \frac{4}{3}\pi a^3 \sigma$. Changing to cylindrical coordinates and

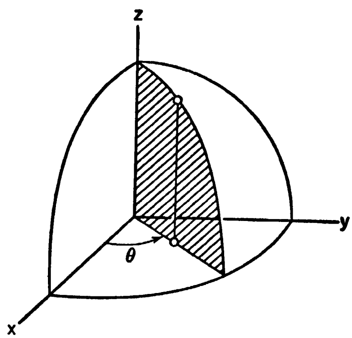



**Fig. 20-29**

setting up the iterated integral, we have $x^2 + y^2 = r^2$, and hence (see Fig. 20-29)

$$I = 2\sigma \int_0^{2\pi} d\theta \int_0^a dr \int_0^{\sqrt{a^2-r^2}} r^3\, dz$$
$$= 2\sigma \int_0^{2\pi} d\theta \int_0^a r^3 \sqrt{a^2 - r^2}\, dr.$$

For the $r$-integration it is convenient to make the substitution $r = a \sin t$, $dr = a \cos t\, dt$. Then (using formula 108 from the table of integrals),

$$\int_0^a r^3 \sqrt{a^2 - r^2}\, dr = a^5 \int_0^{\pi/2} \sin^3 t \cos^2 t\, dt = \frac{2}{3\cdot 5}\, a^5.$$

Hence $I = 4\pi\sigma\frac{2}{15}a^5$. In view of the value of $M$, we can write $I = \frac{2}{5}Ma^2$.

## EXERCISES

1. Write out the limits of integration for the other five orders of integration in the example done in the text. Draw a figure corresponding to Fig. 20-29 for each case.

2. Find the moment of inertia of a homogeneous solid right circular cylinder of radius $b$, about the axis of symmetry.

3. A homogeneous solid right circular cone is defined by $h^2(x^2 + y^2) \leq a^2z^2$, $0 \leq z \leq h$ ($a$ and $h$ positive). Locate the center of mass, using cylindrical coordinates. Set up the limits of integration for the iterated integral in all six orders.

4. Find the moment of inertia of the cone of Exercise 3 about the $z$-axis. Do the problem in two ways, once doing the first integration with respect to $r$, and then first with respect to $z$.

5. Find the attraction of the solid cylindrical shell defined by $a \leq r \leq b$, $0 \leq z \leq h$ (where $0 < a < b$, $0 < h$) on a unit mass particle at the origin, if the solid has constant density.

6. (a) Find the moment of inertia of a homogeneous solid right circular cylinder, of radius $a$ and height $h$, about an axis through the center of mass and perpendicular to the axis of symmetry.
(b) What is the moment of inertia about a diameter of one base of the cylinder?

7. Find the attraction of the cone of Exercise 3 on a unit mass at the origin.

8. The smaller volume cut from the sphere $x^2 + y^2 + z^2 = 4a^2$ by the plane $z = a$ is filled with matter of constant density. Find the attraction which this solid exerts on a unit mass at the origin.

9. Consider the homogeneous solid defined by $(x - a)^2 + y^2 \leq a^2$, $0 \leq z \leq c\sqrt{x^2 + y^2}$. Find the vector representing the attraction which this solid exerts on a unit mass particle at the origin.

## 20-9 Spherical Coordinates

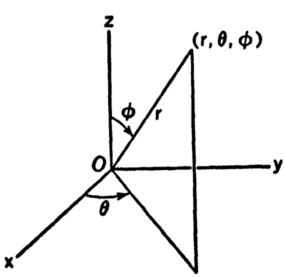


Fig. 20-30

Spherical coordinates (sometimes called spherical polar coordinates) employ the *radial* distance $r$ from $O$ (*not* the $r$ of cylindrical coordinates), the same $\theta$ as in cylindrical coordinates, and an angular coordinate $\phi$ called colatitude. See Fig. 20-30. The relations between rectangular and spherical coordinates are shown in the equations

$$x = r \sin \phi \cos \theta, \qquad y = r \sin \phi \sin \theta, \qquad z = r \cos \phi. \tag{1}$$

Observe that

$$x^2 + y^2 + z^2 = r^2 \quad \text{and} \quad x^2 + y^2 = r^2 \sin^2 \phi.$$

Ordinarily we assume $0 \leq \phi \leq \pi$.

We can calculate a triple integral as an iterated integral in spherical coordinates. To do this we use equations (1) to convert the integrand function $f(x, y, z)$ into a function $F(r, \theta, \phi)$. In order to know what to put in place of $dV$ we think of the process of dividing space into cells of a type appropriate to spherical coordinates. A surface on which $r$ is constant is a sphere with center at $O$. A surface on which $\theta$ is constant is a half-plane through the $z$-axis. A surface on which $\phi$ is constant is a nappe of a cone with vertex $O$ and axis along the $z$-axis. Two surfaces of each type, determined by $r$, $r + \Delta r$, $\theta$, $\theta + \Delta\theta$, $\phi$, $\phi + \Delta\phi$, intersect to form a cell, or "volume element," as shown in Fig. 20-31. The lengths of the three

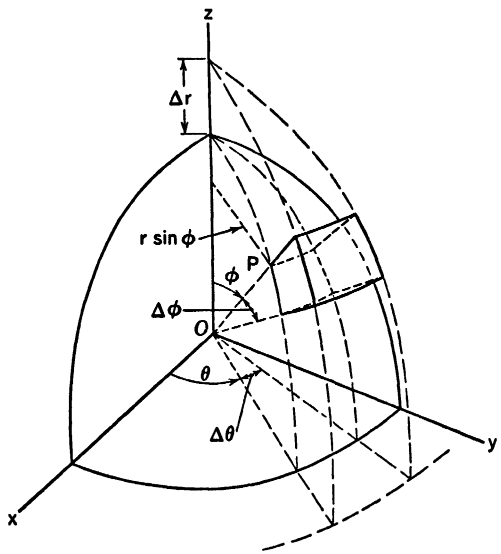



**Fig. 20-31**

mutually perpendicular edges issuing from the point $P(r, \theta, \phi)$ are $\Delta r$, $r\,\Delta\phi$, $r \sin \phi\,\Delta\theta$, so the volume of the cell is approximately $r^2 \sin \phi\,\Delta r\,\Delta\theta\,\Delta\phi$. The *exact* volume of the cell can be shown to be

$$\tfrac{1}{3}[(r + \Delta r)^3 - r^3][\cos \phi - \cos (\phi + \Delta\phi)]\,\Delta\theta. \tag{2}$$

Hence, by application of the law of the mean (Theorem 2-C) separately to the functions $r^3$, $\cos \phi$, we obtain the exact volume of the cell in the form

$$r'^2 \sin \phi'\,\Delta r\,\Delta\theta\,\Delta\phi, \tag{3}$$

where $r'$ is between $r$ and $r + \Delta r$, and $\phi'$ is between $\phi$ and $\phi + \Delta\phi$. It then follows that we are to replace $dV$ by $r^2 \sin\phi \, dr \, d\theta \, d\phi$ and integrate between appropriate limits for the particular region of integration.

**Example 1:** A solid right circular cone is of height $h$ and semivertical angle $\alpha$. If it is filled with matter of uniform density $\sigma$, find the gravitational attraction exerted by the solid on a unit mass particle placed at the vertex.

We place the vertex at the origin and the axis of symmetry along the positive $z$-axis, as shown in Fig. 20-32. By symmetry, the attraction is a force in the positive $z$-direction, of magnitude

$$F = \lambda\sigma \iiint_T \frac{z}{(x^2 + y^2 + z^2)^{3/2}} \, dV,$$

as we see from (3) in § 20-6. Now

$$\frac{z}{(x^2 + y^2 + z^2)^{3/2}} = \frac{\cos\phi}{r^2}.$$

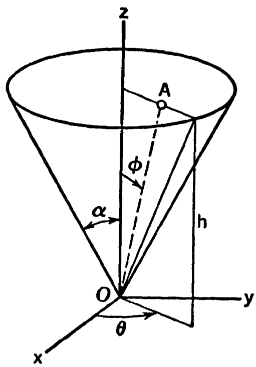



**Fig. 20-32**

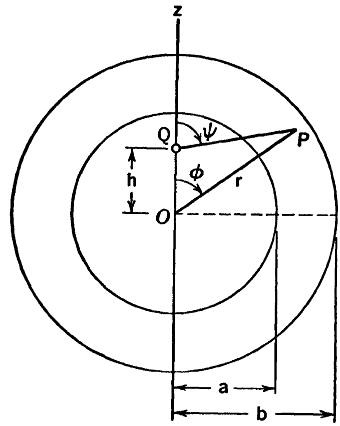



**Fig. 20-33**

Hence, reading the limits of integration from Fig. 20-32, we have

$$F = \lambda\sigma \int_0^{2\pi} d\theta \int_0^{\alpha} d\phi \int_0^{h/\cos\phi} \frac{\cos\phi}{r^2} r^2 \sin\phi \, dr.$$

Note that, for fixed $\theta$ and $\phi$, $r$ varies from $O$ to $OA$, and $OA \cos\phi = h$. After the $r$-integration we have

$$F = \lambda\sigma \int_0^{2\pi} d\theta \int_0^{\alpha} h \sin\phi \, d\phi = 2\pi\lambda\sigma h(1 - \cos\alpha).$$

**Example 2:** Let the space between two concentric spheres be filled with matter of uniform density. Show that the net gravitational attraction of this spherical shell on a unit mass particle is zero if the particle is in the cavity of the shell.

Let the particle be at $Q$, at $z = h$ on the positive $z$-axis, as shown in Fig. 20-33, where $0 < h < a$. Here again the $x$ and $y$ components of attraction are zero, and the component in the $z$-direction is

$$F = \lambda\sigma \iiint_T \frac{\cos\psi}{(PQ)^2}\,dV,$$

where, with $\psi$ as marked in the diagram,

$$(PQ)^2 = r^2 + h^2 - 2hr\cos\phi, \tag{4}$$

$$r\cos\phi - (PQ)\cos\psi = h. \tag{5}$$

From (4) and (5) we see that

$$\frac{\cos\psi}{(PQ)^2} = \frac{r\cos\phi - h}{(r^2 + h^2 - 2rh\cos\phi)^{3/2}}.$$

When our triple integral is converted into an iterated integral, we have

$$F = \lambda\sigma \int_0^{2\pi} d\theta \int_a^b dr \int_0^{\pi} \frac{(r\cos\phi - h)r^2\sin\phi}{(r^2 + h^2 - 2rh\cos\phi)^{3/2}}\,d\phi. \tag{6}$$

To perform the $\phi$-integration we make the substitution

$$u^2 = r^2 + h^2 - 2rh\cos\phi, \qquad u > 0; \tag{7}$$

$$2u\,du = 2rh\sin\phi\,d\phi, \qquad r^2\sin\phi = \frac{ru}{h}\,du.$$

Now, solving for $\cos\phi$ from (7), we find

$$r\cos\phi - h = \frac{r^2 + h^2 - u^2}{2h} - u = \frac{r^2 - h^2 - u^2}{2h};$$

$$\begin{aligned}\int \frac{(r\cos\phi - h)r^2\sin\phi}{(r^2 + h^2 - 2rh\cos\phi)^{3/2}}\,d\phi &= \int \frac{r^2 - h^2 - u^2}{2h}\frac{ru}{hu^3}\,du \\ &= \frac{r}{2h^2}\int\left(\frac{r^2 - h^2}{u^2} - 1\right)du \\ &= \frac{r}{2h^2}\left(-\frac{r^2 - h^2}{u} - u\right) \\ &= -\frac{r}{2h^2}\left(\frac{r^2 - h^2 + u^2}{u}\right).\end{aligned}$$

With $\phi$ as variable this becomes

$$-\frac{r^2(r - h\cos\phi)}{h^2(r^2 + h^2 - 2rh\cos\phi)^{1/2}}.$$

This must be evaluated between the limits $\phi = 0$ and $\phi = \pi$. The result is

$$-\frac{r^2}{h^2}\frac{r + h}{[(r + h)^2]^{1/2}} + \frac{r^2}{h^2}\frac{r - h}{[(r - h)^2]^{1/2}}, \tag{8}$$

which is 0, because for the case we are considering $r - h > 0$ and $[(r - h)^2]^{1/2} = r - h$. Going back now to (6), we see that $F = 0$, since the result of the $\phi$-integration is 0.

## EXERCISES

In these exercises, integrate in spherical coordinates unless directed to do otherwise.

1. Consider the solid sphere $x^2 + y^2 + z^2 \leq a^2$, filled with homogeneous matter.
   (a) Find its moment of inertia about the $z$-axis.
   (b) Locate the center of mass of the hemisphere for which $z \geq 0$.
   (c) Locate the center of mass of the part of the sphere in which $y \geq 0$ and $4(x^2 + y^2 + z^2) \geq a^2$.
   (d) Show that the hemisphere in which $x \geq 0$ attracts a unit mass particle at the origin as though the total mass of the hemisphere were concentrated at the point $(a\sqrt{\frac{2}{3}}, 0, 0)$.

2. Locate the center of mass of the solid cone of the illustrative example in the text (Fig. 20-32).

3. Locate the center of mass of the solid cut from the sphere of Exercise 1 by the nappe of the cone $x^2 + y^2 = z^2 \tan^2 \alpha$ on which $z \geq 0$. Here $0 < \alpha < \pi/2$.

4. Consider the region $T$ inside the sphere $x^2 + y^2 + (z - a)^2 = a^2$ and outside the sphere $x^2 + y^2 + (z - b)^2 = b^2$, where $0 < b < a$.
   (a) Convert these equations to spherical coordinates and set up the limits of integration for integrating over $T$, using spherical coordinates and integrating first with respect to $r$.
   (b) Calculate the mass and locate the center of mass of $T$, thought of as a homogeneous solid. What is the limiting position of the center of mass as $a \to b$?

5. Let the sphere $x^2 + y^2 + (z - a)^2 = a^2$ be filled with homogeneous matter.
   (a) Show that the solid attracts a unit mass particle placed at the origin just as though the mass of the sphere were all concentrated at its center.
   (b) Find the attraction exerted on the unit mass at $O$ by the hemisphere in which $z \geq a$.
   (c) Calculate directly by integration the attraction on the unit mass at $O$ by the hemisphere in which $z \leq a$. Observe that the hemisphere must be divided into two parts for the integration.
   (d) If the density were $\sigma = r$ instead of being constant, show that the entire sphere would attract a unit mass at $O$ as though all the mass of the sphere were concentrated at $(0, 0, \sqrt{\frac{8}{3}}a)$.

6. Find the equation of the sphere $(x - a)^2 + y^2 + z^2 = a^2$ in spherical coordinates. Regarding the hemisphere for which $z \geq 0$ as a homogeneous solid, find the vector attraction of this solid on a particle of unit mass at the origin.

7. (a) If $Q$ were outside the shell in Fig. 20-33 (i.e., if we had $h > b$), show that the attraction of the shell on the unit mass at $Q$ would be of magnitude $\lambda M/h^2$, where $M$ is the mass of the shell. Examine (8) carefully.

(b) Imagine a very small tubular hole bored through a homogeneous solid sphere. Neglecting the effect of this removal of matter, show that the attraction due to the sphere on a mass particle in the tube is directed toward the center of the sphere and of magnitude proportional to the distance the particle is from the center. This can be done by combining the results of Example 2 and Exercise 5(a).

**8.** Derive the expression (2) for the exact volume of the spherical volume element. Start by finding the formula for the volume cut from the sphere $x^2 + y^2 + z^2 = r^2$ by one nappe of the cone $x^2 + y^2 = z^2 \tan^2 \phi$. This volume is that of a right circular cone plus that of a certain spherical segment. The volume of the segment can be found by the method of § 6-7.

**9.** (a) A tetrahedron has its faces in the planes $y = 0$, $z = a$, $x = z$, $x = y$. It is a solid of density inversely proportional to distance from the $z$-axis. Find the mass, using spherical coordinates. Begin by expressing the equations of the planes $z = a$, $x = z$ in spherical coordinates.
(b) Solve the problem using cylindrical coordinates.

# CHAPTER XXI

# DIFFERENTIAL EQUATIONS

## 21-1 Introductory Remarks

The theory of differential equations is a subject of tremendous variety and extent. One can hardly begin it with a complete definition of its scope and purpose. Roughly speaking, in the study of differential equations we attempt to find functions which satisfy certain conditions which are stated in the form of equations involving the unknown function and one or more of its derivatives. The early part of the study is usually made by considering differential equations of a few particularly simple types. Many interesting geometrical and physical problems can be posed as problems in differential equations. Our study in this chapter has three goals: (1) We shall meet and deal with a number of important elementary applications of differential equations. (2) We shall learn a certain amount about the classification of the most readily solvable types of differential equations. (3) We shall progress to the point of understanding how functions not previously known to us become known as the solutions of differential equations.

The simplest of all differential equations is

$$\frac{dy}{dx} = f(x), \tag{1}$$

where $f$ is a specified function of $x$. The problem posed by the equation is that of finding a function $F$ such that if $y = F(x)$, then (1) holds for all values of $x$ in a preassigned interval, or at least for some specified collection of values of $x$. The possibility of finding such an $F$ is affected by the nature of the function $f$ and the specified values of $x$. For example, it is not possible to find such an $F$ for all $x$ such that $0 < x < 2$ if $f$ is the function defined by $f(x) = 0$ when $0 < x < 1$ and $f(x) = 1$ when $1 \leq x < 2$. For a discussion of this see Exercise 1. However, if we assume that $f$ is continuous when $a \leq x \leq b$, then there *does* exist an $F$ defined on $[a, b]$ such that $F'(x) = f(x)$ when $a \leq x \leq b$. One such $F$ is

$$F(x) = \int_a^x f(t)\, dt,$$

and *every* such $F$ is of the form

$$F(x) = \int_a^x f(t)\, dt + C,$$

where $C$ is some constant. These assertions follow from Theorems 6-C and 2-D.

Differential equations occur in interesting and important ways in connection with problems of motion. For instance, we know that if a point moves with simple harmonic motion on the $x$-axis, with the origin as the mid-point of the interval of oscillation, then the $x$-coordinate is a function of $t$ such that

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \tag{2}$$

where $2\pi/\omega$ is the period. This equation (2) is called a second-order differential equation, because of the occurrence of the second derivative.

When Newton's second law of motion is applied to a mass particle moving on the $x$-axis and acted on by certain forces, we obtain a differential equation from which we hope to be able to determine $x$ as a function of $t$. Examples were considered in § 5-6. In a mechanical problem of this sort, the general state of affairs is that

$$\frac{d^2x}{dt^2} = f\left(t, x, \frac{dx}{dt}\right). \tag{3}$$

That is, the acceleration of the particle at a given time $t$ is determined by where the particle is, by its velocity, and by the time itself. In particular cases the function which occurs in (3) may not actually depend on all three of the quantities $t$, $x$, $dx/dt$, but just on one or two of them. The case of constant acceleration is especially simple. It was studied in § 2-3.

Most of the differential equations we study in this chapter are of the first or second order. Let $F$ be an unknown function of $x$. An equation which expresses a condition jointly on $x$, $y$, and $dy/dx$, where $y = F(x)$, is called a *first-order* differential equation. The condition may not involve $x$ or $y$ explicitly, but it must involve $dy/dx$. If the condition is placed on $x$, $y$, $dy/dx$, and $d^2y/dx^2$, with $d^2y/dx^2$ actually involved, the equation is said to be of *second order*. These definitions are made for the purpose of classifying types of differential equations, as a first step in an orderly discussion of how to solve various kinds of differential equations. It is clear how one can go on to define differential equations of third order, fourth order, and so on.

**Examples:**

$$3y + x\frac{dy}{dx} = 9xy^2 \quad \text{(first order);}$$

$$x\frac{d^2y}{dx^2} + (3x^2 + 1)\frac{dy}{dx} = x^3 \quad \text{(second order);}$$

$$y^2 = x\sqrt{1 + \left(\frac{dy}{dx}\right)^2} \quad \text{(first order);}$$

$$\left(\frac{d^2y}{dx^2}\right)^2 = \left[1 + \left(\frac{dy}{dx}\right)^2\right]^3 \quad \text{(second order).}$$

EXERCISES

1. Suppose $f(x) = 0$ if $0 < x < 1$, $f(x) = 1$ if $1 \leq x < 2$. Show that it is impossible to find a function $F$ defined and differentiable for *all* $x$ such that $0 < x < 2$ and such that $F'(x) = f(x)$ for all these values of $x$. Begin by showing what $F(x)$ must be like for $0 < x < 1$ and for $1 < x < 2$, assuming that an $F$ of the required sort does exist. Where does the impossibility show up?

2. (a) Exhibit an integral formula which shows what $F$ must be like if $y = F(x)$ and $d^2y/dx^2 = f(x)$ when $a \leq x \leq b$, given that $f$ is continuous when $a \leq x \leq b$.
(b) Select the particular $F$ which meets the condition of (a) if in addition $F'(a) = A$ and $F(a) = B$.
(c) Select the particular $F$ which meets the condition of (a) if in addition $F(a) = F(b) = 0$.

## 21-2 First-Order Equations with Variables Separable

If $f$ is a function of two variables, the equation

$$\frac{dy}{dx} = f(x, y) \tag{1}$$

is a rather general type of differential equation of the first order. By a solution of this equation we shall mean a differentiable function $F$ of $x$ such that

$$F'(x) = f[x, F(x)] \tag{2}$$

for all $x$ on some interval. There is no a priori reason why such a solution should exist, but one may impose conditions on $f$ which suffice to guarantee the existence of solutions.

We shall proceed by assuming a rather special condition on the function $f$. We assume that $f$ can be expressed as the quotient of a continuous function of $x$ by a continuous function of $y$:

$$\frac{dy}{dx} = \frac{g(x)}{h(y)}. \tag{3}$$

This enables us to write

$$h(y)\,dy = g(x)\,dx. \tag{4}$$

Here we have separated the variables. Once the variables have been separated we can proceed tentatively by forming antiderivatives. An illustration will be given presently.

First-order differential equations often occur in the form

$$M\,dx + N\,dy = 0,$$

where $M$ and $N$ are functions of $x$ and $y$. This equation can be written in the form (1):

$$\frac{dy}{dx} = -\frac{M(x, y)}{N(x, y)},$$

but we need not do this to see whether the variables can be separated.

**Example 1:** Consider the equation

$$x\sqrt{1 - y}\,dx - \sqrt{1 - x^2}\,dy = 0.$$

We separate the variables by writing the equation in the form

$$\frac{x\,dx}{\sqrt{1 - x^2}} = \frac{dy}{\sqrt{1 - y}}. \tag{5}$$

From this form we proceed by taking antiderivatives: Since

$$\int \frac{x\,dx}{\sqrt{1 - x^2}} = -\sqrt{1 - x^2} + C_1$$

and

$$\int \frac{dy}{\sqrt{1 - y}} = -2\sqrt{1 - y} + C_2,$$

we conclude that *if* there is a function $y = F(x)$ satisfying (5), then

$$-\sqrt{1 - x^2} = -2\sqrt{1 - y} + C, \tag{6}$$

where $C$ is some constant. From here we can go on to solve for $y$ and find

$$y = 1 - \frac{1}{4}(\sqrt{1 - x^2} + C)^2. \tag{7}$$

Our work so far shows that if $y = F(x)$ is a solution of (5), then this solution is included in formula (7); that is, for a certain value of $C$, $F$ is given by

$$F(x) = 1 - \frac{1}{4}(\sqrt{1 - x^2} + C)^2$$

on the interval where $F$ is a solution of (5). It still remains to investigate whether (7) does in fact furnish a solution of (5). As we shall see, it *does*, subject to certain conditions. From (7) we obtain

$$(\sqrt{1 - x^2} + C)^2 = 4(1 - y),$$

and thence

$$\sqrt{1 - x^2} + C = \pm 2\sqrt{1 - y}.$$

If the + sign is chosen, we get (6), and from that we get (5). The choice of the + sign is justified if $\sqrt{1 - x^2} + C \geq 0$, but not otherwise. Hence, if $C \geq 0$, (7) defines a solution of the differential equation if $-1 < x < 1$. If $-1 < C < 0$, we get a solution provided that $\sqrt{1 - x^2} \geq -C$, i.e., $|x| \leq \sqrt{1 - C^2}$. But if $C \leq -1$, there is no interval on which (7) is a solution of (5).

The procedure illustrated in Example 1 is useful in practice as a method of seeking solutions of first-order differential equations when the variables can be separated. However, after one gets a relation such as (6) by forming antiderivatives, it may be impractical to solve explicitly for $y$ as we did in going to (7). What one then has, at any rate, is an equation of a one-parameter family of curves in the $xy$-plane. Knowledge of these curves is, in a certain sense, knowledge of solutions of the differential equation, even though one may not have an explicit formula for $y$ in terms of $x$.

**Example 2:** Find a solution of the equation

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

such that $y = 2$ when $x = 4$.

Separating the variables, we have

$$2y\,dy = -x\,dx, \qquad y^2 = -\frac{x^2}{2} + C.$$

In this case we obtain the family of ellipses

$$\frac{x^2}{2} + y^2 = C.$$

The particular one which goes through (4, 2) is the one for which $C = 12$, as we see by substitution. The explicit solution of our problem is

$$y = \sqrt{12 - \frac{x^2}{2}}.$$

*Orthogonal Trajectories*

Suppose we have before us a one-parameter family of smooth curves in the $xy$-plane. By an *orthogonal trajectory* of this family we mean a curve which crosses the curves of the given family at right angles. There are many interesting examples of orthogonal trajectories. For instance, in § 7-4, each circle of the family (1) is an orthogonal trajectory of the family of circles given by (2). In the case of confocal families of ellipses and hyperbolas (in § 7-5), the hyperbolas are orthogonal trajectories of the ellipses, and the ellipses are orthogonal trajectories of the hyperbolas.

The finding of orthogonal trajectories of a given family of curves involves a two-stage problem: (1) From the equation of the given family eliminate the parameter by differentiation and thus obtain the slope at $(x, y)$ as a function of $x$ and $y$, say

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

(2) For the orthogonal trajectory through $(x, y)$ the slope is the negative reciprocal of the former slope, so we have

$$\frac{dy}{dx} = -\frac{1}{f(x, y)} \tag{8}$$

as the slope for the orthogonal trajectory. The family of curves obtained by solving (8) will be the orthogonal trajectories of the original family.

**Example 3:** Find the orthogonal trajectories of the family of parabolas (with parameter $k$):

$$x^2 = 2ky. \tag{9}$$

As the first stage of the solution we eliminate $k$:

$$2x\,dx = 2k\,dy \quad \text{and} \quad k = \frac{x^2}{2y},$$

so

$$\frac{dy}{dx} = \frac{x}{k} = \frac{2y}{x}.$$

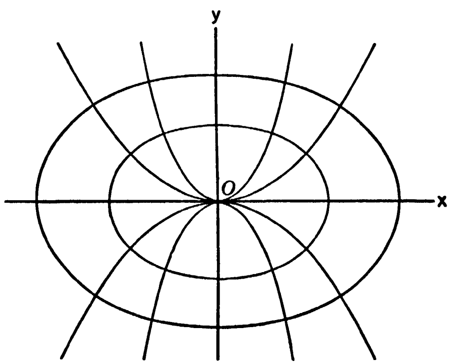



**Fig. 21-1**

The second stage of the solution begins when we write

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

as the differential equation of the orthogonal trajectories. This equation is the one which was considered in Example 2; as we saw there, it leads us to the family of ellipses

$$\frac{x^2}{2} + y^2 = C. \tag{10}$$

The characteristic feature of the parabolas is that they have foci on the $y$-axis and vertices at the origin. The ellipses have foci on the $x$-axis, centers at the origin, and they all have the same eccentricity, $e = \sqrt{2}/2$. Two curves of each type are shown in Fig. 21-1.

## EXERCISES

**1.** Find a one-parameter family of curves which includes every solution of the given differential equation, in the sense explained in Example 1.

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

(b) $(y + 5)\,dx = (x + 3)\,dy.$

(c) $\dfrac{dy}{dx} = -\dfrac{\sin x \cos^2 y}{\cos^2 x}.$

(d) $3\sqrt{a^2 - y^2}\,dx + y\,dy = 0.$

(e) $\dfrac{dy}{dx} = \dfrac{2y \sin x \cos x}{\sin^2 x - \cos^2 x}.$

**2.** Proceed as directed in Exercise 1.

(a) $\dfrac{dy}{dx} = \dfrac{xy}{\sqrt{1 - x^2}}.$

(b) $x\dfrac{dy}{dx} = y^2.$

(c) $\dfrac{dy}{dx} = e^{3x-7y}.$

(d) $xy\,dx - \dfrac{1 + x^2}{1 + y^2}\,dy = 0.$

(e) $2x\sqrt{y^2 + 4}\,dx + y\sqrt{9x^2 - 16}\,dy = 0.$

**3.** Find the solution of $\dfrac{dy}{dx} = \dfrac{y}{x^2 - 4}$ such that

(a) $y = 2$ when $x = 8$;

(b) $y = 2$ when $x = 2$;

(c) $y = -2$ when $x = 0$.

**4.** Find the solution of $\dfrac{dy}{dx} = \sqrt{\dfrac{1 - y^2}{1 - x^2}}$ such that

(a) $y = 1/\sqrt{2}$ when $x = \frac{1}{2}$;

(b) $y = \frac{4}{5}$ when $x = \frac{3}{5}$.

5. (a) Find a one-parameter family of curves each of which has a slope at $(x, y)$ such that $\sin x\,dx - \cos x \cos y\,dy = 0$.
(b) Find the solution of the differential equation such that $y = \pi/6$ when $x = 0$.

6. A family of curves in the first quadrant has the property that the straight line tangent to one of the curves at $(x, y)$ cuts the $x$-axis at $(3x, 0)$.
(a) Find the equation of the family.
(b) Find the particular curve through the point $(4, 1)$.

7. A family of curves in the first quadrant has the property that the straight line tangent to one of the curves at $(x, y)$ cuts the $y$-axis at $(0, xy)$.
(a) Find the equation of the family.
(b) Find the particular curve through the point $(1, e)$.

8. Find the orthogonal trajectories of the given family of curves. Draw a few of the given curves and a few of the trajectories. Use differential equations even if the solution is geometrically obvious.
(a) $x^2 + y^2 = k$.
(b) $x^2 - y^2 = k$.
(c) $ky = x^3$.
(d) $y^2 = kx^3$.

9. Proceed as directed in Exercise 8.
(a) $y = k \sin x,\ 0 < x < \pi/2$.
(b) $y = \dfrac{k}{x^2 + 1}$.
(c) $y = ke^{-x^2}$.
(d) $y - \sqrt{1 - x^2} = k$.

## 21-3 First-Order Equations and One-Parameter Families

The purpose of this section is to orient the student with respect to the geometrical meaning of first-order differential equations and their solutions. No systematic methods for solving problems are presented in this section, and the exposition is descriptive and intuitive rather than formal, analytical, and precise or complete.

Consider the first-order equation

$$\frac{dy}{dx} = f(x, y), \tag{1}$$

where $f$ is assumed to be continuous in a region of the $xy$-plane. For simplicity let us suppose that the region is a rectangle $R$ with each side parallel to a coordinate axis. We can then imagine what is called a *direction-field* in $R$. Through each point $(x, y)$ of $R$ we draw a short line segment having as its slope the value of $f$ at $(x, y)$. This assemblage of line segments is the visual representation of the direction-field. See Fig. 21-2. One of the line segments is called a *direction-element* of the field. We say that the direction-element constructed through the point $(x, y)$ is *asso-*

*ciated with the point.* Now, if there is a curve $y = F(x)$ in $R$ such that at each of its points it is tangent to the direction-element associated with the point, it is clear that $F'(x) = f[x, F(x)]$, and hence that $y = F(x)$ is a solution of the differential equation (1). The curve $y = F(x)$ is then called an *integral curve* of the differential equation.

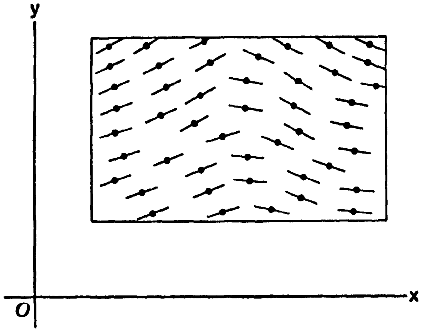



Fig. 21-2

Suppose now that we have a one-parameter family of smooth curves coursing through a rectangle $R$, as shown in Fig. 21-3. We suppose that through any given point of $R$ there passes exactly one curve of the family, and that no curve ever has a tangent (at a point in $R$) parallel to the $y$-axis. Then, at each point $(x, y)$ in $R$ there is a unique slope of the curve through that point. This defines a function $f$: $f(x, y)$ = slope at $(x, y)$ of the curve which passes through $(x, y)$. Then each curve of the family is an integral curve of the differential equation $y' = f(x, y)$, and a direction-field can be constructed by drawing segments of lines tangent to the curves of the family.

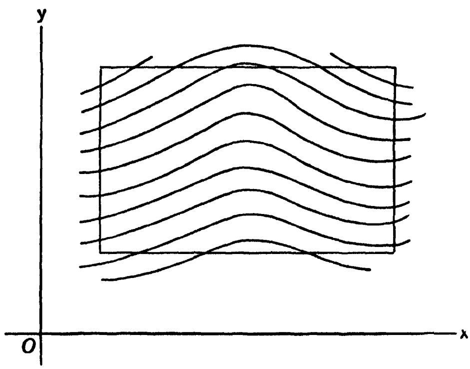



Fig. 21-3

If we merely have the differential equation $y' = f(x, y)$, and we imagine the corresponding direction-field to have been constructed, it is natural to speculate as to whether there really does exist a family of curves having the direction-elements as tangents. In the theory of differential equations, at a more advanced level than our present one, it is shown that, if certain assumptions are made about $f$, one can prove the existence of a unique family of curves giving rise to the specified direction-field. It is sufficient to assume that $f$ and $\partial f/\partial y$ are continuous in $R$.

Let us now suppose there is such a unique family of integral curves. If we fix our attention on the integral curves which pass through or near some one point $(x_0, y_0)$ in $R$, we see that they form a one-parameter family. For example, let $y_1$ be any value of $y$ near $y_0$. Then there is a unique integral curve through $(x_0, y_1)$. By varying $y_1$ (regarding $y_1$ as a parameter) we get the one-parameter family of all integral curves passing through or near $(x_0, y_0)$.

In many comparatively simple problems it is possible to start with a one-parameter family of smooth curves and find a first-order differential equation of which they are all integral curves, the method of procedure being to eliminate the parameter by differentiation. This procedure was illustrated in the discussion of orthogonal trajectories in § 21-2. Further illustration was provided in Exercises 8, 9 of § 21-2. A careful theoretical discussion of this procedure would involve the use of implicit function theorems. The process may not always be practical in the sense of elementary algebra.

**Example 1:** Consider the family of parabolas

$$2py + p^2 = x^2, \tag{2}$$

with $p$ as the parameter, assuming $p > 0$.

We differentiate with respect to $x$:

$$2p\frac{dy}{dx} = 2x, \qquad \frac{dy}{dx} = \frac{x}{p}. \tag{3}$$

Now we solve (2) for $p$:

$$p = \frac{-2y \pm \sqrt{4y^2 + 4x^2}}{2} = -y \pm \sqrt{y^2 + x^2}.$$

Since we are assuming $p > 0$, we must choose the $+$ sign. When the proper value of $p$ is put into (3) we obtain

$$\frac{dy}{dx} = \frac{x}{\sqrt{x^2 + y^2} - y} \tag{4}$$

as the differential equation of the family of parabolas.

**Example 2:** Consider the family of curves $y = Ce^{Cx}$.

Here the elimination of $C$ in an elementary explicit way is not possible. Conceptually we proceed as follows: Differentiation gives

$$\frac{dy}{dx} = C^2e^{Cx} = Cy. \tag{5}$$

If $x$ and $y$ are positive, a graphical discussion of the equation

$$C = ye^{-Cx}$$

shows that it is satisfied by a unique positive value of $C$, which is thereby determined as a function of $x$ and $y$, say $C = g(x, y)$. Putting this in (5), we obtain

$$\frac{dy}{dx} = yg(x, y)$$

as the differential equation of the family, in so far as it lies in the first quadrant.

Many of the elementary procedures for solving differential equations of particular types lead ultimately to the formation of antiderivatives. When this occurs, an arbitrary constant is brought in. With first-order equations there is just one arbitrary constant brought in in this way, and we obtain a one-parameter family of integral curves. The mere fact of having a one-parameter family does not always guarantee that from this family one can obtain all the integral curves passing near a particular point. For example, $y = e^{x+C}$ is a one-parameter family of solutions of $dy/dx = y$; it yields some, but not all, the integral curves which pass near the origin. The trouble is that if we write $y = e^C e^x$, the constant $e^C$ cannot be negative or zero. A more comprehensive one-parameter family is $y = Ce^x$, where $C$ may be assigned any real value. This family *does* yield all solutions of the given differential equation; this may be shown by the methods of § 21-5.

## 21-4 Homogeneous First-Order Equations

When a first-order equation is written in the form

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

it may happen that $M$ and $N$ are homogeneous functions of the same degree. The differential equation is then called *homogeneous*. This means that there is some index $p$ (not necessarily an integer) such that

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

for all suitably restricted values of $x$, $y$, $t$, and likewise for $N$. The index $p$ is the *degree*.

**Example 1:** The equation

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

is homogeneous. The degree is 2, and there are no restrictions on $x$, $y$, $t$.

**Example 2:** The equation

$$(y - \sqrt{x^2 + y^2})\, dx - x\, dy = 0$$

is homogeneous. The degree is 1 and the restriction on $t$ is that $t \geq 0$.

When $M$ and $N$ in (1) are homogeneous of the same degree, there is a device whereby with a change of variable we can convert the differential equation into a new form and separate the variables. The device consists in taking either $y/x$ or $x/y$ as a new variable. If $y/x = v$, the new variables are taken to be $v$ and $x$. If $x/y = u$, we use $y$ and $u$ as new variables. The homogeneity comes in as follows: If $y = vx$, then, at least for suitable values of the variables,

$$M(x, y) = M(x, vx) = x^p M(1, v).$$

We treat $N$ in the same way. Also,

$$dy = v\,dx + x\,dv.$$

It then turns out that the variables can be separated.

**Example 3:** Consider the equation

$$(x - 2y)\,dx + y\,dy = 0. \tag{3}$$

With $y = vx$ we get

$$(x - 2vx)\,dx + vx(v\,dx + x\,dv) = 0,$$

$$x(1 - 2v + v^2)\,dx + x^2 v\,dv = 0,$$

$$\frac{dx}{x} + \frac{v\,dv}{(v-1)^2} = 0.$$

The variables are now separated, and we can proceed. To calculate one of the antiderivatives it is convenient to let $v - 1 = t$. In this way we obtain

$$\log|x| + \log|v - 1| - \frac{1}{v-1} = \text{constant}. \tag{4}$$

It is convenient to denote the constant by $\log C$, where $C > 0$. Then

$$\log\frac{|x(v-1)|}{C} = \frac{1}{v-1}.$$

Putting $v = y/x$, we find

$$\log\frac{|y - x|}{C} = \frac{x}{y - x}. \tag{5}$$

There are many interesting problems which lead to differential equations of the homogeneous type. A homogeneous differential equation can be written in the form

$$\frac{dy}{dx} = g\left(\frac{y}{x}\right), \tag{6}$$

where $g$ is a function of one variable. From this it appears that, in the direction-field associated with the equation (6), all the direction-elements at points along a line $y = mx$ have the same slope, namely, $g(m)$.

## EXERCISES

1. (a) Find the family of curves determined by the differential equation $(x^2 + y^2)\,dx - 2xy\,dy = 0$. Identify the curves by name and draw a number of them.

(b) Find the orthogonal trajectories of the curves in (a). Draw the first quadrant portion of some of the curves. Observe that you can solve for $x$ in terms of $y$.

2. Find the orthogonal trajectories of the family of circles $(x - k)^2 + y^2 = k^2$. A proper use of symmetry will yield great dividends.

3. Find families of integral curves for the following differential equations:
(a) $(y^2 - xy)\,dx + x^2\,dy = 0$.
(b) $(y^2 - x^2 + xy)\,dx + (y^2 - x^2 - 2xy)\,dy = 0$.
(c) $(4x - y)\,dx + (x + y)\,dy = 0$.

4. Follow the directions of Exercise 3.
(a) $\dfrac{dy}{dx} = \dfrac{2x + y}{x + 2y}$.
(b) $(y\sqrt{x^2 - y^2} - xy)\,dx + x^2\,dy = 0$.
(c) $(x^2 + y^2)\,dx + 2y(x + y)\,dy = 0$.

5. (a) Show that the introduction of polar coordinates transforms the differential equation

$$\frac{dy}{dx} = f\left(\frac{y}{x}\right)$$

into one in which the variables can be separated, and that the family of integral curves can be expressed in the form

$$r = Ce^{F(\theta)},$$

where $F(\theta)$ is a certain antiderivative.
(b) Apply the method of (a) to the equation

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

(c) Apply the method of (a) to the equation

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

(d) Solve the problem in (c) by the method explained in the text, and compare with the result obtained in polar coordinates.

## 21-5 The General First-Order Linear Equation

We continue with our program of studying the equation $y' = f(x, y)$ in various cases where $f(x, y)$ has an especially tractable form. One very important case is that in which $f(x, y)$ is a linear function of $y$, with coefficients depending on $x$. That is, we consider

$$f(x, y) = p(x)y + q(x),$$

where $p$ and $q$ are continuous functions of $x$ on some interval of the $x$-axis. It turns out to be more convenient to introduce a minus sign and write

$$f(x, y) = -P(x)y + Q(x),$$

so that the differential equation is

$$\frac{dy}{dx} + P(x)y = Q(x). \tag{1}$$

This is the general form of what is called a *first-order linear* differential equation.

There is a simple explicit procedure by which the equation (1) can be solved. Let $\phi$ be an antiderivative of $P$. That is, let $\phi$ be a function, defined on the same interval as $P$ and $Q$, such that

$$\phi'(x) = P(x). \tag{2}$$

Then $$\frac{d}{dx}[ye^{\phi(x)}] = ye^{\phi(x)}\phi'(x) + \frac{dy}{dx}e^{\phi(x)}.$$

Hence, in view of (2),

$$e^{\phi(x)}\left[\frac{dy}{dx} + P(x)y\right] = \frac{d}{dx}[ye^{\phi(x)}]. \tag{3}$$

Now, since $e^{\phi(x)}$ is never zero, the equation (1) is equivalent to

$$e^{\phi(x)}\left[\frac{dy}{dx} + P(x)y\right] = Q(x)e^{\phi(x)},$$

which, because of (3), can be written

$$\frac{d}{dx}[ye^{\phi(x)}] = Q(x)e^{\phi(x)}. \tag{4}$$

Thus, $y = F(x)$ is a solution of (1) if and only if $F(x)e^{\phi(x)}$ is an antiderivative of $Q(x)e^{\phi(x)}$. It follows that

$$y = e^{-\phi(x)}\int Q(x)e^{\phi(x)}\,dx \tag{5}$$

is a solution of (1), and that every solution of (1) can be exhibited in this form.

**Example:** Find a general representation of the solutions of

$$(y - \sin x)\cos x\,dx + \sin x\,dy = 0.$$

This equation can be brought into the standard first-order linear form:

$$\frac{dy}{dx} + y \operatorname{ctn} x = \cos x. \tag{6}$$

Here $P(x) = \operatorname{ctn} x$; an antiderivative of it is $\log \sin x$, provided that $\sin x > 0$. If $\sin x < 0$ we can take $\log(-\sin x)$ as the antiderivative. It turns out at the end that we get the same final formula for our solution in either case. Then

$$e^{\log \sin x} = \sin x.$$

We utilize the method by which (5) was derived, rather than trying to remember (5) itself. That is, we multiply (6) through by $\sin x$, and then we have

$$\sin x \frac{dy}{dx} + y \cos x = \frac{d}{dx}(y \sin x) = \sin x \cos x.$$

Consequently

$$y \sin x = \int \sin x \cos x \, dx = \frac{1}{2} \sin^2 x + C,$$

and

$$y = \frac{1}{2} \sin x + C \csc x. \tag{7}$$

Because (7) is a solution of (6) for every choice of the value of $C$, and because every solution of (6), on an interval where $\csc x$ is continuous, is included in the solutions given by (7), we call (7) the *general* solution of (6).

The structure of the solution (7) can be analyzed in the following way: The solution is composed by addition of the two parts $\frac{1}{2} \sin x$ and $C \csc x$. The part $\frac{1}{2} \sin x$ is a solution of (6):

$$\frac{d}{dx}(\tfrac{1}{2} \sin x) + (\tfrac{1}{2} \sin x) \operatorname{ctn} x = \cos x.$$

The part $C \csc x$, with the arbitrary coefficient $C$, *is not* a solution of (6), but it *is* a solution of the differential equation

$$\frac{dy}{dx} + y \operatorname{ctn} x = 0, \tag{8}$$

which is obtained from (6) when the right member of (6) is replaced by 0:

$$\frac{d}{dx}(C \csc x) + (C \csc x) \operatorname{ctn} x = 0.$$

This same kind of analysis can be made in the case of the general solution of (1). If $G(x)$ is a particular antiderivative of $Q(x)e^{\phi(x)}$, so that

$$G'(x) = Q(x)e^{\phi(x)}, \tag{9}$$

then formula (5) in its generality can be written in the form

$$y = e^{-\phi(x)}G(x) + Ce^{-\phi(x)},$$

where $C$ is an arbitrary constant. Here the first part, $e^{-\phi(x)}G(x)$, is a solution of (1), because of (2) and (9); the second part, $Ce^{-\phi(x)}$, is the general solution of the equation

$$\frac{dy}{dx} + P(x)y = 0.$$

This analysis of the structure of the general solution of (1) is given here primarily to provide a background of understanding for a similar situation in relation to linear differential equations of the second order. Such equations are discussed in § 21-8.

## EXERCISES

**1.** Find the general solution of each equation.

(a) $\dfrac{dy}{dx} + xy = x.$

(b) $\dfrac{dy}{dx} - y \operatorname{ctn} x = \sin x.$

(c) $x\dfrac{dy}{dx} - y = x \log x.$

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

(e) $x^2\dfrac{dy}{dx} + (1 - 2x)y = x^2.$

**2.** Proceed as directed in Exercise 1.

(a) $(x - y + 1)\, dx = x\, dy.$

(b) $2x\, dy = (x^3 - x + y)\, dx.$

(c) $\sin x \dfrac{dy}{dx} - y \cos x = x \sin^2 x.$

(d) $x(x^2 - y)\, dx = dy.$

(e) $(1 - x^2)^{3/2}\, dy + y\, dx = dx.$

**3.** Consider a circuit (say a coil of wire) in which an electric current is flowing. The physical law governing the current is expressed by the equation

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

Here $L$ and $R$ are positive constants: $L$ is the self-inductance of the circuit and $R$ is the resistance of the circuit. The current strength is denoted by $i$ and the applied electromotive force is $E$; both $i$ and $E$ are to be regarded as functions of the time $t$, though $E$ may, as a particular case, be constant. Express the current as a function of $t$ in each of the following cases, assuming that $i = i_0$ when $t = 0$. Observe that the current in each case is expressible as the sum of two parts: a *transient part,* which depends on $i_0$ and approaches 0 as $t \to +\infty$, and a *steady-state part,* which is independent of $i_0$.

(a) $E$ is a constant.

(b) $E = E_0 \sin \omega t$ ($E_0$ and $\omega$ constant).

**4.** Given the situation of Exercise 3, find $i$ as a function of $t$ in each of the following cases, assuming that $i = 0$ when $t = 0$.

(a) $E = E_0 e^{-Rt/L}.$

(b) $E = E_0 e^{-Rt/L} \cos \omega t.$

**5.** If we have an electric circuit with a constant resistor $R$, an electromotive force $E$, and a capacitor of capacitance $C$ arranged in series, the charge $q$

on the capacitor satisfies the equation

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

This equation determines $q$ as a function of $t$ and the initial value of $q$ ($q = q_0$ when $t = 0$). The current in the circuit is then given by $i = dq/dt$. Solve for $q$:

(a) If $E$ is constant.

(b) If $E = E_0 \cos \omega t$.

6. A differential equation of the form

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

where $n \neq 1$, can be converted into a first-order linear equation by the change of variable $u = y^{1-n}$, with $u$ as the new dependent variable. This form of equation is called *Bernoulli's equation*. The form occurs in some interesting mechanical problems.

Demonstrate the validity of the assertion made about converting the given equation to the first-order linear form. Then apply the method to the following particular cases, and find the general solution in each case.

(a) $\frac{dy}{dx} - \frac{y}{2} = xy^{-3}$.

(b) $\cos\theta \frac{dv}{d\theta} + v \sin\theta + v^2 = 0$.

(c) $6y\frac{dy}{dx} = 18 \sin x - 3 \cos x + y^2$.

## 21-6 Miscellaneous Applications

In this section we shall discuss several concrete problems whose solutions illustrate in an interesting way the uses of differential equations.

**Example 1:** Two points $A$, $O$ are directly opposite each other on the banks of a river of width $a$. A man starts at $A$ and rows across the river, always heading directly toward $O$. If the river current is uniform, and if the man's rate of rowing in still water is equal to the speed of the current, find the curve described by the boat.

We choose axes as indicated in Fig. 21-4. If $v$ is the speed of the river current, the man's components of velocity are

$$\frac{dx}{dt} = v - v\cos\theta, \qquad \frac{dy}{dt} = -v\sin\theta.$$

Therefore $\frac{dy}{dx} = \frac{-v\sin\theta}{v - v\cos\theta} = \frac{\sin\theta}{\cos\theta - 1}$.

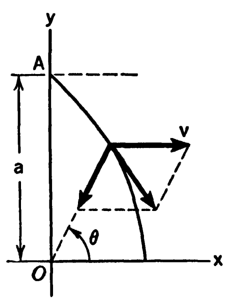



Fig. 21-4

But $$\sin\theta = \frac{y}{\sqrt{x^2 + y^2}}, \qquad \cos\theta = \frac{x}{\sqrt{x^2 + y^2}}.$$

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

This is a homogeneous differential equation. It is convenient to take $x/y = u$ as a new variable (though we could take $y/x$ instead). Then we have

$$(yu - \sqrt{y^2u^2 + y^2})\,dy = y(y\,du + u\,dy),$$

from which we find

$$\frac{dy}{y} + \frac{du}{\sqrt{u^2 + 1}} = 0.$$

Then $$\log y + \log (u + \sqrt{u^2 + 1}) = C,$$

or $$\log (yu + \sqrt{y^2u^2 + y^2}) = C.$$

We now put back $x = yu$ and use the fact that $y = a$ when $x = 0$, thus finding that $C = \log a$. Therefore

$$x + \sqrt{x^2 + y^2} = a.$$

If we free this equation of radicals, it takes the form

$$y^2 = -2a\left(x - \frac{a}{2}\right),$$

which is the equation of a parabola with vertex at $(a/2, 0)$. This shows that the man will reach the bank of the river at a point downstream from $O$ at a distance from $O$ equal to half the width of the river.

**Example 2:** A surface of revolution has the property that the volume bounded by the surface and two planes perpendicular to the axis of revolution is directly proportional to the area of the part of the surface between the two planes. Find the shape of the surface.

A right circular cylinder obviously has the required property. We therefore dismiss it and seek other surfaces. Let us suppose that the surface is generated by revolving a curve $y = f(x)$ about the $x$-axis. The volume of the solid of revolution between the planes determined by $x_1$ and $x$ is

$$V = \pi \int_{x_1}^{x} [f(t)]^2\,dt.$$

The corresponding area on the surface of revolution is

$$S = 2\pi \int_{x_1}^{x} f(t)\sqrt{1 + [f'(t)]^2}\,dt.$$

These formulas come from § 6-1 and § 11-4, respectively. But $V$ is proportional to $S$; that is, $V = kS$. Hence

$$\frac{dV}{dx} = k\frac{dS}{dx}.$$

But $$\frac{dV}{dx} = \pi[f(x)]^2, \qquad \frac{dS}{dx} = 2\pi f(x)\sqrt{1 + [f'(x)]^2}.$$

Therefore, with $y = f(x)$, we have

$$y^2 = 2ky\sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \tag{1}$$

Our problem now is to solve this differential equation. It is convenient to let $c = 2k$. Then, if $y \neq 0$, (1) implies that

$$\left(\frac{dy}{dx}\right)^2 = \frac{y^2 - c^2}{c^2}. \tag{2}$$

One obvious solution is $y = f(x) \equiv c$. This is the case of the right circular cylinder. Apart from this solution we have

$$\frac{dy}{\sqrt{y^2 - c^2}} = \pm\frac{dx}{c}.$$

Passing to antiderivatives and using formula (6) of § 9-3, we get

$$\cosh^{-1}\frac{y}{c} = \pm\frac{x}{c} + A,$$

where $A$ is the constant of integration. Then

$$y = c\cosh\left(\pm\frac{x}{c} + A\right).$$

If we locate our axes in such a way that $y = c$ when $x = 0$, then $A = 0$ and

$$y = c\cosh\frac{x}{c} = \frac{c}{2}(e^{x/c} + e^{-x/c}). \tag{3}$$

The double sign disappears because the hyperbolic cosine is an even function. The curve is called a *catenary*.

**Example 3:** Consider a situation in which a heavy bead is sliding on a rough circular hoop of wire, the hoop standing in a vertical plane, as shown in Fig. 21-5. Suppose $\mu$ is the coefficient of friction. Let $R$ be the normal reaction of the hoop on the bead. Suppose the bead starts at $\theta = 0$ with a small initial value of $d\theta/dt$. Find $d\theta/dt$ subsequently as a function of $t$.

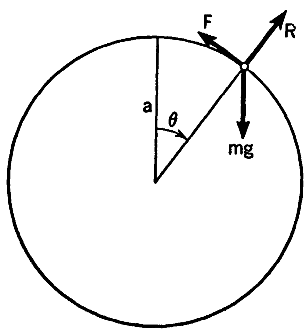



**Fig. 21-5**

The force of friction is denoted by $F$. It is a force tangent to the hoop, directed in the sense opposite to that in which the bead moves. Its magnitude is $F = \mu|R|$. In the early stage of the motion $R$ is directed outward, and we regard $R$ as positive when it is directed outward. Then $F = \mu R$. We use Newton's law and resolve the acceleration into tangential and radial components. From an analysis of tangential forces and accelerations we conclude that

$$ma\frac{d^2\theta}{dt^2} = mg\sin\theta - F.$$

From the radial components we obtain the equation

$$ma\left(\frac{d\theta}{dt}\right)^2 = mg\cos\theta - R.$$

Hence, since $F = \mu R$,

$$ma\frac{d^2\theta}{dt^2} = mg\sin\theta - \mu\left[mg\cos\theta - ma\left(\frac{d\theta}{dt}\right)^2\right],$$

or

$$\frac{d^2\theta}{dt^2} - \mu\left(\frac{d\theta}{dt}\right)^2 = \frac{g}{a}(\sin\theta - \mu\cos\theta). \tag{4}$$

This is a second-order differential equation for the determination of $\theta$, but we can turn it into a first-order equation for the determination of $d\theta/dt$. The device for doing this is as follows: Let $p = d\theta/dt$. Then

$$\frac{d^2\theta}{dt^2} = \frac{dp}{dt} = \frac{dp}{d\theta}\frac{d\theta}{dt} = p\frac{dp}{d\theta}.$$

In this way (4) becomes

$$p\frac{dp}{d\theta} - \mu p^2 = \frac{g}{a}(\sin\theta - \mu\cos\theta). \tag{5}$$

This is an equation of Bernoulli type. Discussion of the solution of it is left for the Exercises.

## EXERCISES

**1.** Suppose the man in Example 1 of the text can row with twice the speed of the current.

(a) Show that his path is part of the curve $4ax^2 = y(a - y)^2$.

(b) Where does he reach the opposite bank, and in what direction is he headed when he reaches the bank?

(c) How long does it take him to cross the river? (To answer this, express $dy/dt$ in terms of $y$, using the result of (a). Then integrate from $y = a$ to $y = 0$.)

**2.** If the man in Example 1 of the text can row with half the speed of the current, show that his path is part of the curve $2a^2xy = a^4 - y^4$. Does he ever get across?

**3.** The speed of the current of a river is proportional to the product of the distances to the two banks.

(a) If, as in Example 1 of the text, a man starts at $A$, but always rows directly toward the opposite bank, how far downstream will he be carried before he gets across? Assume the man's rate of rowing is constant, and equal to the speed of the current in midstream.

(b) What is the equation of his path?

4. When water issues from an orifice in a container, its velocity is $\sqrt{2gh}$, where $h$ is the vertical distance from the orifice up to the water surface in the container, and $g$ is the acceleration due to gravity. The effective rate of efflux from a small, sharp-edged orifice of area $A$ is approximately $0.6A\sqrt{2gh}$ cubic units per second. The factor 0.6 is accounted for by friction and a certain contraction in the size of the stream.
(a) Find the time required to empty a cylindrical container through a hole 1 inch in diameter in the bottom. Suppose the radius of the cylinder is 1 foot, and let the water be initially 3 feet deep.
(b) Suppose, with the container as in (a), water is also running into the container at the rate of $\pi/2$ cubic feet per minute. Show that the depth of the water will never get as small as 1 foot, but that it will approach this figure as a limit. If the water was initially 4 feet deep, how long will it take the depth to decrease to 2 feet?

5. A curve $y = f(x)$ passes through the origin and goes into the first quadrant. At each point $P$ of the curve in the first quadrant lines are drawn parallel to the axes, thus forming a rectangle with diagonally opposite corners at $O$ and $P$. It is then found that the area inside the rectangle and under the curve is one third of the area of the rectangle. This being true for each position of $P$, find the equation of the curve.

6. A sphere has the property that the area of any zone is directly proportional to the distance between the bases of the zone. Show that, aside from circular cylinders, the sphere is the only surface of revolution that has this property.

7. What curve $y = f(x)$, $[f(x) > 0]$ has the property that the area between the curve and the $x$-axis, from $x = x_1$ to $x = x_2$, is in constant ratio to the arc length of the curve between these same values of $x$?

8. Find the general solution of equation (5) in Example 3 of the text.

9. (a) If we suppose that the population $y$ of the United States increases according to the law

$$\frac{dy}{dt} = ky(300 - y),$$

$y$ being measured in millions and $t$ in decades, obtain a formula for $y$ in terms of $t$. Use the data $y = 76$ in 1900 ($t = 0$), $y = 92$ in 1910 ($t = 1$), to determine the constant $k$ and the constant of integration.
(b) What is the result given by the formula for 1980?
(c) With these assumptions what is the limiting value of the population?

10. A particle of mass $m$ is set in motion under water. It then moves in a vertical plane, acted on by the force of gravity and a drag due to resistance by the water. Assume that this resistance is a vector of magnitude $kmv$, opposite in direction to the vector velocity. Here $k$ is a constant and $v = ds/dt$ is the speed, $s$ being measured along the curve in the direction

of the motion. See Fig. 21-6. Show that the relation between $v$ and $\phi$ is given by the differential equation

$$\frac{1}{v}\frac{dv}{d\phi} = \frac{kv}{g}\sec\phi + \tan\phi.$$

Hence, show that

$$v = \frac{g\sec\phi}{Cg - k\tan\phi},$$

where $C$ is a constant depending on the initial conditions. You will need to use results from Chapter XIII, especially as regards tangential and normal components of acceleration.

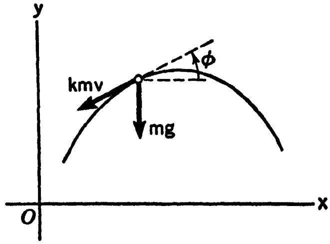



**Fig. 21-6**

**11.** (a) On the assumption that air resistance is proportional to the square of the speed, the velocity of an object of mass $m$ falling freely through the atmosphere is governed by the equation $m\frac{dv}{dt} = mg - kv^2$, where $k$ is a constant ($v$ positive for downward motion). Solve this equation, assuming initially $v = 0$, $t = 0$. Show that when $t$ is very large, $v$ is approximately $\sqrt{mg/k}$.
(b) In a scientific test a man jumped from an airplane and fell 29,300 feet before opening his parachute. His total weight, with equipment, was 285 pounds. Instruments showed that he reached a limiting velocity of 230 miles per hour. Using the differential equation from (a), find the value of $k$, and calculate the number of seconds required for the man to attain 99 per cent of his limiting velocity.

**12.** The velocity of a small lead shot of mass $m$ falling vertically through water obeys the law

$$\frac{dv}{dt} + \frac{a}{m}v = \left(1 - \frac{1}{\rho}\right)g,$$

where $a = 1.69 \times 10^{-2}$, $g = 980$, and $\rho$ is the density of lead. Units are those of the cgs. system. Calculate the limiting velocity of the shot, and the time required to attain half this velocity, starting from rest. Assume $\rho = 11$ and consider the shot to be a sphere of radius 0.05 centimeter.

**13.** (a) Consider a flexible cord or chain hanging over a horizontal circular cylinder of radius $b$, with $\mu$ the coefficient of friction between the cord and the cylinder. Let $\sigma$ be the linear density (constant) of the cord. Use a diagram somewhat like Fig. 8-9 in § 8-6, but take into account the weight, $g\sigma b\,\Delta\theta$, of the segment of the cord corresponding to $\Delta\theta$. Let $\theta = 0$ be the horizontal direction. If the cord is on the point of slipping in the direction of increasing $\theta$, show that the tension $T$ is determined by the differential equation

$$\frac{dT}{d\theta} - \mu T = \sigma g b(\mu\sin\theta + \cos\theta).$$

(b) Solve this equation, assuming that one end of the cord is at $\theta = 0$ (and hence $T = 0$ there), while the other end hangs down a distance $h$ below the point $\theta = \pi$ on the other side of the cylinder. What is the value of $h$?

14. A dome has the shape of a surface of revolution with the following property: if a stone is dropped from the window of a nearby building, at exactly the level of the top of the dome, the horizontal projection of the stone on the dome will move along the dome with constant speed. Show that the profile of the dome is a cycloid.

15. A snowfall begins at some time in the forenoon. It snows steadily on into the afternoon. At noon a man begins to clear the sidewalk on a certain street. He shovels two blocks by 2 o'clock and one block more by 4 o'clock. At what time did the snow begin to fall? Assume that the man removes snow at a fixed number of cubic feet per hour, and that the sidewalk is of uniform width. The man does not go back to clear the snow that has fallen behind him.

## 21-7 Equations of the Second Order. Some Special Types

In this section we consider differential equations of the second order of two special types. In the general case, if $y$ is dependent and $x$ is independent, a second-order equation may involve all four of the quantities

$$x, \quad y, \quad \frac{dy}{dx}, \quad \frac{d^2y}{dx^2}.$$

However, if either $x$ or $y$ does not occur explicitly, we can deal with the equation by converting it to an equation of the first order.

*Dependent Variable Absent*

In this case we can let

$$p = \frac{dy}{dx}, \qquad \frac{dp}{dx} = \frac{d^2y}{dx^2}. \tag{1}$$

Then the differential equation is of first order in the variables $x$, $p$. If this first-order equation can be solved, the solution will provide us with a first-order equation in the variables $x$, $y$, and we can then address ourselves to the solution of the latter equation.

**Example 1:** The equation

$$x\frac{d^2y}{dx^2} - 2\frac{dy}{dx} = x^3$$

becomes

$$x\frac{dp}{dx} - 2p = x^3,$$

which is of the first-order linear type.

Solving as in § 21-5, we find

$$p = \frac{dy}{dx} = x^3 + C_1x^2.$$

(The student should supply the details.) We can now integrate directly:

$$y = \frac{1}{4}x^4 + \frac{1}{3}C_1x^3 + C_2.$$

The solution involves two arbitrary constants.

*Independent Variable Absent*

In this case we again let $p = dy/dx$. But now we write

$$\frac{d^2y}{dx^2} = \frac{dp}{dx} = \frac{dp}{dy}\frac{dy}{dx} = p\frac{dp}{dy}. \tag{2}$$

This is the same as the device that was used in (2) of § 5-6. By means of it our second-order equation becomes a first-order equation in $y$ and $p$. If we can solve it, the solution furnishes us a differential equation of the first order in $x$ and $y$.

**Example 2:** We saw in § 20-4 that the motion of a compound pendulum is governed by the differential equation

$$\frac{d^2\theta}{dt^2} = -\frac{Mgh}{I}\sin\theta, \tag{3}$$

where $h$ is the distance from the point of support $O$ to the center of mass of the pendulum, $M$ is the mass, and $I$ is the moment of inertia about a horizontal axis through $O$ (see Fig. 21-7). We shall write $l = I/Mh$, so that (3) becomes

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\sin\theta. \tag{4}$$

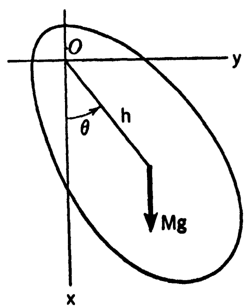


Fig. 21-7

Let us try to solve (4), assuming $\theta = \theta_0$ and $d\theta/dt = 0$ when $t = 0$. The value of $\theta_0$ is taken so that $0 < \theta_0 < \pi$. Here the roles of $x$, $y$ are taken by $t$, $\theta$. Instead of (2) we have

$$p = \frac{d\theta}{dt}, \qquad \frac{d^2\theta}{dt^2} = p\frac{dp}{d\theta}.$$

Then (4) becomes $$p\frac{dp}{d\theta} = -\frac{g}{l}\sin\theta.$$

The variables can be separated, and we obtain

$$\frac{p^2}{2} = \frac{g}{l}\cos\theta + C.$$

Since $p = 0$ when $\theta = \theta_0$, we can calculate $C$. Then

$$\left(\frac{d\theta}{dt}\right)^2 = p^2 = \frac{2g}{l}(\cos\theta - \cos\theta_0).$$

At the next step, when we extract a square root, we must decide which sign to take. For the time while $\theta$ is decreasing we have

$$\sqrt{\frac{l}{2g}}\frac{d\theta}{\sqrt{\cos\theta - \cos\theta_0}} = -dt.$$

If $T$ is the period (the time for one complete oscillation), the time for $\theta$ to decrease from $\theta_0$ to 0 is $T/4$. Hence

$$\sqrt{\frac{l}{2g}}\int_{\theta_0}^{0}\frac{d\theta}{\sqrt{\cos\theta - \cos\theta_0}} = -\int_0^{T/4} dt = -\frac{T}{4}. \tag{5}$$

The integral on the left is an elliptic integral unless $\cos\theta_0 = -1$. It can be changed into a more standard form by writing

$$\cos\theta = 1 - 2\sin^2\frac{\theta}{2}$$

and introducing a new variable $\phi$ by the relation

$$\sin\frac{\theta}{2} = k\sin\phi, \qquad \text{where } k = \sin\frac{\theta_0}{2}. \tag{6}$$

This leads to the formula

$$T = 4\sqrt{\frac{l}{g}}\int_0^{\pi/2}\frac{d\phi}{\sqrt{1 - k^2\sin^2\phi}}. \tag{7}$$

Details are left to the student. The integral in (7) is called an *elliptic integral of the first kind*. There are tables giving the value of the integral as a function of $k$.

For quite small values of $\theta$ (up to about the radian equivalent of 5°) it is a satisfactory approximation to replace $\sin\theta$ by $\theta$ in the differential equation (4). The new differential equation is that which is characteristic of simple harmonic motion. In other words, when the amplitude of oscillation of the pendulum is sufficiently small, the angle $\theta$ can be expressed approximately in the form

$$\theta = \theta_0\cos\sqrt{\frac{g}{l}}\,t$$

(when the initial conditions are those given here). For larger amplitudes this is not a valid approximation. It is nevertheless true that $\theta$ is a periodic function of $t$. It is, however, a more complicated type of function. It is called an *elliptic* function.

*Loaded Flexible Cables*

Consider a flexible, inextensible cable which is loaded with a weight of amount $w$ per unit length of arc (so that the total downward force on a part of the cable is given by the integral

$$\int_{s_1}^{s_2} w\, ds$$

extended over the part in question). Here $w$ may be variable from point to point along the cable. We suppose the cable to be strung between two supports; our problem is to find the curve in which the cable hangs when in equilibrium. See Fig. 21-8. We have chosen our axes so that $dy/dx = 0$ at $x = 0$ (the lowest point of the cable).

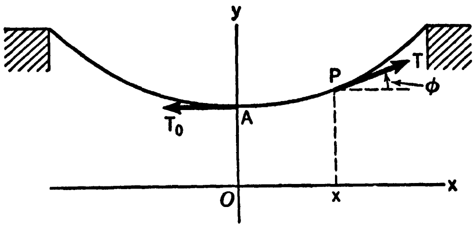



**Fig. 21-8**

The conditions of equilibrium for the section of the cable from $A$ to $P$ require that the tension $T$ at $P$ satisfy the equations

$$T \cos \phi = T_0 \quad \text{and} \quad T \sin \phi = \int w\, ds,$$

where the integration is along the arc from $A$ to $P$. Then $d(T \sin \phi) = w\, ds$. But

$$T \sin \phi = \frac{T_0}{\cos \phi} \sin \phi = T_0 \tan \phi = T_0 \frac{dy}{dx},$$

and so we have

$$\frac{d}{dx}\left(T_0 \frac{dy}{dx}\right) = w \frac{ds}{dx},$$

or

$$T_0 \frac{d^2y}{dx^2} = w \frac{ds}{dx}. \tag{8}$$

To go further with this differential equation we must know something definite about $w$.

**Example 3:** Suppose the only load carried by the cable is that of its own weight. Then $w$ is constant. To proceed with (8) we have

$$\frac{d^2y}{dx^2} = \frac{w}{T_0} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \tag{9}$$

Using (1) and setting $h = T_0/w$ for convenience, we have

$$\frac{dp}{dx} = \frac{1}{h}\sqrt{1 + p^2}, \qquad \frac{dp}{\sqrt{1 + p^2}} = \frac{dx}{h}.$$

Forming antiderivatives by (5) in § 9-3, we have

$$\sinh^{-1} p = \frac{x}{h} + C_1, \qquad p = \frac{dy}{dx} = \sinh\left(\frac{x}{h} + C_1\right).$$

Since $p = 0$ when $x = 0$, we find that $C_1 = 0$. We can now integrate once more:

$$y = h \cosh\frac{x}{h} + C_2.$$

We recognize this as the equation of a catenary [see (3) in § 21-6]. If we locate the axes so that $y = h$ when $x = 0$, then $C_2 = 0$. In this position the $x$-axis is called the *directrix* of the catenary.

## EXERCISES

**1.** Solve each differential equation, obtaining a solution with two arbitrary constants. If proper data are given, use them to evaluate the constants.

(a) $x\dfrac{d^2y}{dx^2} = \dfrac{dy}{dx}$.

(b) $\left(\dfrac{d^2y}{dx^2}\right)^2 = \left[1 + \left(\dfrac{dy}{dx}\right)^2\right]^3$; $x = 0$, $y = 1$, $\dfrac{dy}{dx} = 0$.

(c) $1 + \left(\dfrac{dy}{dx}\right)^2 + y\dfrac{d^2y}{dx^2} = 0$; $x = 1$, $y = 1$, $\dfrac{dy}{dx} = 0$.

(d) $x\dfrac{d^2y}{dx^2} - \dfrac{dy}{dx}\left[1 + \left(\dfrac{dy}{dx}\right)^2\right] = 0$.

**2.** Follow the directions of Exercise 1.

(a) $(a^2 - x^2)\dfrac{d^2y}{dx^2} = a\sqrt{1 + \left(\dfrac{dy}{dx}\right)^2}$; $x = 0$, $y = -a$, $\dfrac{dy}{dx} = 0$.

(b) $\left(\dfrac{d^2y}{dx^2}\right)^2 = \left[1 - \left(\dfrac{dy}{dx}\right)^2\right]^3$; $x = 0$, $y = 1$, $\dfrac{dy}{dx} = 0$.

(c) $y\dfrac{d^2y}{dx^2} = 1 + \left(\dfrac{dy}{dx}\right)^2$.

(d) $1 + \left(\dfrac{dy}{dx}\right)^2 + 2y\dfrac{d^2y}{dx^2} = 0$.

**3.** Suppose the load on the cable in Fig. 21-8 is that of a horizontal roadway of constant weight $c$ per unit length in the $x$-direction. Neglect the weight of everything except the roadway itself. Show that the curve of the cable is $y = (cx^2/2T_0) + h$, where $h = OA$.

**4.** Find the shape assumed by the cable in Fig. 21-8 if the load on any length of the cable is $k$ times the area directly underneath that length and above some horizontal line. Choose the horizontal line as the $x$-axis, and denote the value of $y$ when $x = 0$ by $h$.

**5.** For the catenary in Example 3 above, show that the tension is $T = T_0 \cosh(x/a)$. Verify that this is just the weight of a section of the cable long enough to extend from $P$ down to the directrix.

**6.** It may be shown by the principles of mechanics for rigid bodies, as explained in § 20-4, that a ladder of length $l$ with one end on a smooth floor and the other end against a smooth wall, will slide down according to the law expressed in the differential equation

$$\frac{d^2\theta}{dt^2} = \frac{3g}{l}\sin\theta,$$

where $\theta$ is the angle the ladder makes with the wall. Assume that $\theta = \theta_0$ and $d\theta/dt = 0$ when $t = 0$ and deduce that the general relation between $\theta$ and $t$ is

$$t = \sqrt{\frac{l}{3g}}\int_{\theta_0}^{\theta}\frac{d\theta}{\sqrt{\cos\theta_0 - \cos\theta}}.$$

By letting $\cos(\theta/2) = k\sin\phi$, $k = \cos(\theta_0/2)$, convert this result to the form

$$t = \sqrt{\frac{2l}{3g}}\int_{\phi}^{\pi/2}\frac{d\phi}{\sqrt{1 - k^2\sin^2\phi}}.$$

**7.** Consider the cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$. With the axes as shown in Fig. 21-9, suppose that a heavy bead is sliding on the cycloid,

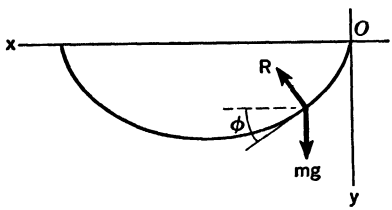



**Fig. 21-9**

thought of as a smooth wire. Show that the time required for the bead to slide to the lowest point is $\pi\sqrt{a/g}$, regardless of where the bead starts to slide, provided it has no initial velocity. The solution of this problem falls into two parts. First, use Newton's law, taking into account the tangential component of acceleration, and get the equation $dv/dt = g(dy/ds)$, where $v = ds/dt$, $s$ being measured from $O$. From this deduce that $v^2 = 2g(y - y_0)$, where $y_0$ denotes the initial value of $y$. Then use the relation between $s$ and $\theta$ on the cycloid to show that the time for the bead to reach the lowest point is

$$t = \sqrt{\frac{2a}{g}}\int_{\theta_0}^{\pi}\frac{\sin(\theta/2)}{\sqrt{\cos\theta_0 - \cos\theta}}\,d\theta.$$

The problem can then be finished by a device like that used in the last part of Exercise 6.

8. From Exercise 11(a) in § 21-6 show that the distance $s$ fallen in time $t$ is $\frac{m}{k}\log\cosh\sqrt{\frac{gk}{m}}\,t$. Show also that

$$v^2 = \frac{mg}{k}(1 - e^{-2ks/m}).$$

## 21-8 Linear Equations of the Second Order

The general form of a linear differential equation of the second order is

$$A(x)\frac{d^2y}{dx^2} + B(x)\frac{dy}{dx} + C(x)y = D(x). \tag{1}$$

We shall assume that the functions $A$, $B$, $C$, $D$ are continuous on some interval of the $x$-axis. If $D$ is the zero function, the differential equation is said to be homogeneous in $y$. (This is a different use of the word homogeneous from that in § 21-4.) In order to abbreviate our notation conveniently, we shall write

$$L[F] = A(x)F''(x) + B(x)F'(x) + C(x)F(x). \tag{2}$$

If $y = F(x)$, we shall sometimes write $L[y]$ instead of $L[F]$. Observe that, if $c$ is a constant,

$$L[cF] = cL[F]. \tag{3}$$

Also, if $F$ and $G$ are two functions,

$$L[F + G] = L[F] + L[G]. \tag{4}$$

The simple facts expressed in (3) and (4) are very important for what we are now going to discuss. Notice that with our notation the differential equation (1) can be written in the form

$$L[y] = D(x). \tag{5}$$

We say that $y = F(x)$ is a solution of this equation on a given interval if $L[F] = D(x)$ for each $x$ on the interval.

In the study of the problem of solving the second-order linear differential equation (5) it turns out to be important to study also the problem of solving the equation

$$L[y] = 0. \tag{6}$$

As we have said already, the equation (6) is called homogeneous. *We call* (6) *the homogeneous equation corresponding to equation* (5). Sometimes we call (6) the *reduced* equation instead of the *homogeneous* equation. Several simple facts can be noted. They are so useful that we state them as theorems

THEOREM 21-A. *If* $y = F_1(x)$ *and* $y = F_2(x)$ *are solutions of the homogeneous equation* (6), *then every linear combination with constant coefficients,* $y = c_1F_1(x) + c_2F_2(x)$, *is also a solution.*

This is an immediate consequence of (3) and (4).

THEOREM 21-B. *If* $F_1$ *and* $F_2$ *are two solutions of* (5), *the difference* $F_2 - F_1$ *is a solution of the corresponding reduced equation* (6).

*Proof.* We assume that $L[F_1] = L[F_2] = D(x)$. Then, by (3) and (4),

$$L[F_2 - F_1] = L[F_2] - L[F_1] = D(x) - D(x) = 0.$$

The next very important thing to consider is the concept of the *general solution* of equation (5), or of (6). When we speak about "the general solution" of $L[y] = D(x)$ for a certain interval of the $x$-axis, we mean a family of solutions which includes all possible solutions.

For our purposes we shall assume that the coefficient $A(x)$ in (1) is never zero on the interval under consideration. This assumption is needed in the theory which justifies some of the assertions we are going to make.

Two functions $F_1$, $F_2$ are said to be *linearly independent* on an interval if neither function is on that interval a constant multiple of the other function. (In particular, this guarantees that neither function can be identically zero.)

THEOREM 21-C. *If* $F_1$ *and* $F_2$ *are linearly independent solutions of* $L[y] = 0$ *on a given interval, and if* $A(x)$ *is never zero on this interval, then the family* $c_1F_1 + c_2F_2$, *where* $c_1$ *and* $c_2$ *are arbitrary constants, is the general solution of* $L[y] = 0$.

The proof of this theorem is beyond the scope of this book. The proof is given in standard texts on the theory of differential equations. We shall use the theorem without attempting to present a proof. The general theory also makes it possible to show, under the given conditions, that there *do exist* two linearly independent solutions.

With the aid of Theorems 21-B and 21-C it is possible to describe a procedure for finding all solutions of the non-reduced equation (5). *First, try to find two linearly independent solutions of the reduced equation* (6). Suppose we do find two such solutions, $F_1$ and $F_2$. *Next, try to find at least one solution of the non-reduced equation* (5). Suppose we are able to find one such solution, say $y = F(x)$. *Then the family*

$$y = c_1F_1(x) + c_2F_2(x) + F(x) \tag{7}$$

*is the general solution of* (5). That is, if $y = G(x)$ is any solution whatsoever of (5), the constants $c_1$, $c_2$ in (7) can be assigned values in such a way that

$$G(x) = c_1F_1(x) + c_2F_2(x) + F(x).$$

The proof is very simple. Since $G$ and $F$ are solutions of (5), $G - F$

is a solution of (6), by Theorem 21-B. But then, by Theorem 21-C, $G - F$ is included in the family $c_1F_1 + c_2F_2$, and hence $G$ is included in the family $c_1F_1 + c_2F_2 + F$.

**Example 1:** Consider the equation

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} - y = 2. \tag{8}$$

The corresponding homogeneous equation is

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} - y = 0. \tag{9}$$

It happens to be rather easy to spot a solution of (8) in this case. If we try to find a constant solution $y = k$, we see that we must have $-k = 2$, or $k = -2$. [Note, incidentally, that (5) will always admit a certain constant solution if $C(x)$ and $D(x)$ are constant functions.] Next, because of the particular structure of the left member of (9), it seems reasonable to attempt to find a power of $x$ which is a solution. Trying $y = x^n$ in (9), we see that we need to have

$$n(n-1)x^n + nx^n - x^n = x^n[n^2 - 1] = 0.$$

This works out fine if $n = \pm 1$. Hence $y = x$ and $y = 1/x$ are solutions of the reduced equation. By what we have learned, then, the general solution of (9) is

$$y = c_1x + \frac{c_2}{x},$$

and the general solution of (8) is

$$y = c_1x + \frac{c_2}{x} - 2.$$

Spotting solutions by guesswork or shrewd observation is in fact a procedure of great usefulness in work with differential equations.

*The Orbit of a Planet*

We shall now use what we have learned to help with the derivation of Kepler's first law from Newton's laws. We start with results worked out in § 13-6.

Consider a planet as a particle moving around the sun as a center of attraction. We ignore all other masses. Let the sun be at $O$ and let the planet have polar coordinates $(r, \theta)$. Then the planet is attracted toward $O$ by a force of magnitude $c/r^2$, where $c$ is a constant. Hence, in view of the formula in § 13-6 for the radial component of acceleration,

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = -\frac{c}{r^2}. \tag{10}$$

We also know from § 13-6 that

$$r^2\frac{d\theta}{dt} = h \quad \text{(a constant)}. \tag{11}$$

By using these two formulas we can obtain a second-order differential equation governing $r$ as a function of $\theta$. It turns out to be more convenient to let $u = 1/r$ and work with $u$:

$$\frac{dr}{dt} = -\frac{1}{u^2}\frac{du}{dt} = -\frac{1}{u^2}\frac{du}{d\theta}\frac{d\theta}{dt}$$

$$= -\frac{1}{u^2}\frac{du}{d\theta}\frac{h}{r^2} = -h\frac{du}{d\theta}.$$

$$\frac{d}{dt}\left(\frac{dr}{dt}\right) = \frac{d}{d\theta}\left(-h\frac{du}{d\theta}\right)\frac{d\theta}{dt} = -h^2u^2\frac{d^2u}{d\theta^2}.$$

When these calculations are combined with (10) and (11), we obtain

$$\frac{d^2u}{d\theta^2} + u = \frac{c}{h^2}. \tag{12}$$

One particular solution of this is obviously $u = c/h^2$. The reduced equation,

$$\frac{d^2u}{d\theta^2} + u = 0,$$

is just like the equation governing simple harmonic motion with period $2\pi$. It has the solutions $u = \sin\theta$, $u = \cos\theta$. Therefore the general solution of (12) is

$$u = C_1 \sin\theta + C_2 \cos\theta + \frac{c}{h^2}.$$

This can also be written in the form

$$u = \frac{c}{h^2} - B\cos(\theta - \theta_0),$$

with $B$ and $\theta_0$ related to $C_1$ and $C_2$ by the equations

$$B\cos\theta_0 = -C_2, \qquad B\sin\theta_0 = -C_1.$$

Going back now to $r$, we have

$$r = \frac{h^2/c}{1 - (Bh^2/c)\cos(\theta - \theta_0)}. \tag{13}$$

We see from § 12-2 that this is the polar form of the equation of either a parabola, an ellipse, or a hyperbola. Thus we see that the planetary orbit must be one of these curves with the sun at a focus. This is Kepler's first law.

*Solution in Series*

If $P$, $Q$, $R$ are functions of $x$ which can be represented by power series in $x$ in some interval containing $x = 0$, it can be proved that the differential equation

$$\frac{d^2y}{dx^2} + P(x)\frac{dy}{dx} + Q(x)y = R(x)$$

has solutions on this interval which can be represented by power series. If $y = F(x)$ is a solution in series form, by substituting directly into the equation one can see how to compute the coefficients of the series for $F$ in terms of the coefficients of the series for $P$, $Q$, $R$. The study of solutions by series is an important part of the theory of differential equations.

### EXERCISES

**1.** The equation $(1 - x^2)y'' - 2xy' + 2y = 0$ has $y = x$ as one solution. Write $y = vx$ and take $v$ as a new dependent variable. In this way obtain

$$y = -1 - \frac{x}{2}\log\frac{1-x}{1+x}$$

as another solution when $-1 < x < 1$. Hence find the general solution of

$$(1 - x^2)y'' - 2xy' + 2y = 6.$$

**2.** Use (11) and (13) to show that if the planetary orbit is an ellipse of semimajor axis $a$, and if the planet goes once around the orbit in time $T$, then $cT^2 = 4\pi a^3$. This is Kepler's third law. You can compute $a$ and $b$ for the ellipse from (7) in § 12-2.

## 21-9 Linear Differential Equations with Constant Coefficients

In this section we concentrate attention on the homogeneous second-order linear equation

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0 \tag{1}$$

in which the coefficients $a$, $b$ are constant. To find the general solution of this equation we know from § 21-8 that it suffices to find two linearly independent solutions. If we take note of the fact that $e^{mx}$ reproduces itself when differentiated, it is natural to attempt to find a solution of (1) by trying $y = e^{mx}$. For this to satisfy (1) we must have

$$e^{mx}(m^2 + am + b) = 0.$$

This method works fine if the quadratic equation

$$m^2 + am + b = 0 \tag{2}$$

has roots which are real and distinct. We call this equation the *auxiliary equation.*

**Example 1:** Consider the equation

$$\frac{d^2y}{dx^2} + \frac{dy}{dx} - 2y = 0.$$

In this case (2) becomes

$$m^2 + m - 2 = (m - 1)(m + 2) = 0,$$

with roots 1, $-2$. Hence the general solution of our differential equation is

$$y = C_1e^x + C_2e^{-2x}.$$

But what if the two roots of (2) are the same, say $m_1 = m_2 = r$? Then this method does not give us two linearly independent solutions, but only one solution, namely, $e^{rx}$. However, in this case it will be found that $xe^{rx}$ is also a solution, and then the general solution is

$$y = C_1e^{rx} + C_2xe^{rx}.$$

Verification is left as an exercise.

If the quadratic equation (2) does not have real roots, we are up against a difficulty in trying to find a solution of the differential equation by this method. Let us illustrate by an example. If we try the method on the equation

$$\frac{d^2y}{dx^2} + 4y = 0,$$

we substitute $y = e^{mx}$ and obtain

$$e^{mx}(m^2 + 4) = 0.$$

This indicates that $m$ should be taken to be $\pm 2i$, so that $e^{2ix}$ and $e^{-2ix}$ should be solutions. But we have not even defined a meaning for $e^{\pm 2ix}$, and without such a definition we can hardly consider it proper to regard $e^{2ix}$ and $e^{-2ix}$ as solutions.

What is needed, evidently, is an adequate definition of $e^{mx}$ when $m$ is a complex number. We also need to know about differentiation of $e^{mx}$ with respect to $x$ when $m$ is complex and $x$ is real. Both of these needs can be attended to, and when they are, the $e^{mx}$ method enables us to find solutions of the differential equation even when the roots of the auxiliary equation are not real.

A complex number $w$ is expressed in terms of a pair of real numbers: $w = u + iv$, where $u$ and $v$ are real. We assume the student knows at least the rudiments of formal algebra for complex numbers. The definition of $e^w$ which we use is that expressed in the following formula:

$$e^{u+iv} = e^u(\cos v + i \sin v), \tag{3}$$

where $e^u$, $\cos v$, $\sin v$ have the meanings already familiar to us. In this brief presentation we shall not attempt to write down a motivation for the definition. Once the definition is given, it is not hard to deduce from it the facts which we need. The foremost of these facts is that

$$\frac{d}{dx} e^{mx} = me^{mx}, \tag{4}$$

even when $m$ is any *complex* constant. The meaning of differentiation of complex-valued functions is simple. If $f_1$ and $f_2$ are real-valued differentiable functions, the derivative of $f_1 + if_2$ is given by

$$\frac{d}{dx}[f_1(x) + if_2(x)] = f_1'(x) + if_2'(x).$$

This may be proved by the $\Delta$-process, just as we proved the rule for sums in § 3-2. The proof of (4) is given as an exercise.

**Example 2:** The differential equation

$$\frac{d^2y}{dx^2} - 4\frac{dy}{dx} + 5y = 0 \tag{5}$$

has the auxiliary equation

$$m^2 - 4m + 5 = 0,$$

with roots

$$m = \frac{4 \pm \sqrt{16 - 20}}{2} = 2 \pm i.$$

Hence $$e^{(2+i)x} = e^{2x}(\cos x + i \sin x)$$

and $$e^{(2-i)x} = e^{2x}(\cos x - i \sin x)$$

are solutions of the differential equation. These are nonreal complex solutions. For some purposes it is desirable to have *real* solutions. Now the coefficients in equation (5) are real numbers. In this case, following the general rule which we shall state in a moment, both the real and the imaginary parts of a complex solution are themselves solutions. Hence, in the present case,

$$e^{2x} \cos x \quad \text{and} \quad e^{2x} \sin x$$

are real solutions of (5). From them we can build the general solution, for they are linearly independent.

The principle used here can be stated as follows:

THEOREM 21-D. *If a homogeneous linear differential equation with real coefficients admits a complex function as a solution, then the real and imaginary parts of this function are also solutions.*

*Proof.* This principle applies to the general case considered in § 21-8, not merely to the case of equations with constant coefficients. If $F_1$ and $F_2$ are real functions, and if the coefficients in the differential equation are real, then

$$L[F_1 + iF_2] = L[F_1] + iL[F_2].$$

Moreover, $L[F_1]$ and $L[F_2]$ are real. Hence $L[F_1 + iF_2] = 0$ implies $L[F_1] = L[F_2] = 0$; this is the proof of the theorem.

### EXERCISES

**1.** Find the general real solution of each differential equation. If some conditions are indicated, find the particular solution which satisfies these conditions.
(a) $y'' - y' - 6y = 0$; $y = 2$ and $y' = -9$ when $x = 0$.
(b) $y'' - 6y' + 9y = 0$; $y = -1$ and $y' = 0$ when $x = 0$.
(c) $y'' + 9y = 0$; $y = 2$ and $y' = 6$ when $x = 0$.
(d) $y'' - 2y' + 2y = 0$; $y = 0$ and $y' = \sqrt{2}$ when $x = \pi/4$.
(e) $y'' + 2y' = 0$.
(*f*) $y'' + 4y' + 13y = 0$.

**2.** Follow the directions of Exercise 1.
(a) $y'' - 3y' - 10y = 0$; $y = 2$ and $y' = 38$ if $x = 0$.
(b) $y'' + 4y' + 4y = 0$; $y = 4$ and $y' = 6$ if $x = 0$.
(c) $y'' + 17y' + 16y = 0$.
(d) $2y'' - 2y' + 13y = 0$.
(e) $y'' - 5y' = 0$; $y = -3$ and $y' = 2$ when $x = 0$.
(*f*) $y'' + 2y' + 10y = 0$.

**3.** If the auxiliary equation (2) has a double root $r$, show that (1) takes the form $y'' - 2ry' + r^2y = 0$. Show that $e^{rx}$ and $xe^{rx}$ are solutions of this equation.

**4.** Suppose that $m = p + iq$, where $p$ and $q$ are real. Use (3) to express $e^{(p+iq)x}$ in terms of its real and imaginary parts. Then calculate the derivative with respect to $x$ and show that it is the same as $(p + iq)e^{(p+iq)x}$, when calculated by the usual rule for multiplying complex numbers. This constitutes a proof of (4).

**5.** Find the general solution of the equation

$$(L^2 - M^2)\frac{d^2I}{dt^2} + 2RL\frac{dI}{dt} + R^2I = 0.$$

## 2-10 Oscillatory Systems

The simplest oscillatory motion is simple harmonic motion. It is characterized by the differential equation

$$\frac{d^2x}{dt^2} + \omega^2x = 0,$$

whose general solution can be written in either of the forms

$$x = A\cos\omega t + B\sin\omega t \quad (A, B \text{ arbitrary}),$$

$$x = C\cos(\omega t - \alpha) \quad (C, \alpha \text{ arbitrary}).$$

Pendulum motion is oscillatory, but its differential equation is nonlinear.

*Damping*

A differential equation of the form

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + a^2x = 0, \tag{1}$$

where $a$ and $\lambda$ are positive, arises when a system which would otherwise undergo simple harmonic motion is subjected to a retarding force which is proportional to the speed. This retarding force accounts for the term involving $dx/dt$ in (1). The size of $\lambda$ is an indication of the damping effect of the retarding force. There are three cases to consider.

*Case I.* $a > \lambda$. Here there is relatively little damping. We set $n = \sqrt{a^2 - \lambda^2}$. The roots of the auxiliary equation are $-\lambda \pm in$, and the general solution of the differential equation is

$$x = e^{-\lambda t}(A \cos nt + B \sin nt).$$

The motion is oscillatory in character, with frequency $n$, but the factor $e^{-\lambda t}$ causes the amplitude to approach zero as $t \to \infty$.

*Case II.* $a < \lambda$. In this case the damping is severe. If we set $p = \sqrt{\lambda^2 - a^2}$, the general solution is

$$x = e^{-\lambda t}(Ae^{pt} + Be^{-pt}).$$

There is no oscillation. Since $0 < p < \lambda$, we see that $x \to 0$ as $t \to \infty$. As an example, consider the effect of a very stiff hydraulic door-check on a swinging door.

*Case III.* $a = \lambda$. In this case the auxiliary equation has equal roots. The general solution of the differential equation is

$$x = (A + Bt)e^{-\lambda t}.$$

Again there is no oscillation, but $x$ may increase in absolute value for a time before tending toward zero.

*Forced Vibrations*

If a force of oscillatory character is impressed on the free system, we get what are called *forced vibrations*. A typical case would be that of the differential equation

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + a^2x = E \cos pt, \tag{2}$$

where $E$ and $p$ are constants. Since we know the solution of the reduced equation, the problem is to find *one* solution of (2). We attempt to find such a solution as a linear combination of cos $pt$ and sin $pt$:

$$x = C_1 \cos pt + C_2 \sin pt. \tag{3}$$

On substituting this into (2) and collecting like terms we have

$$\begin{pmatrix} -C_1p^2 \\ +2\lambda C_2 p \\ +a^2C_1 \end{pmatrix} \cos pt + \begin{pmatrix} -C_2p^2 \\ -2\lambda C_1 p \\ +a^2C_2 \end{pmatrix} \sin pt = E \cos pt.$$

Thus, in order for (3) to be a solution of (2), it is sufficient that

$$(a^2 - p^2)C_1 + 2\lambda pC_2 = E,$$
$$-2\lambda pC_1 + (a^2 - p^2)C_2 = 0.$$

We can solve this pair of equations uniquely for $C_1$ and $C_2$ if

$$(a^2 - p^2)^2 + 4\lambda^2p^2 \neq 0.$$

Thus, barring an exceptional case, we see that the differential equation (2) does admit a solution of the form (3). This solution has the same frequency as the impressed force, but its amplitude is different and there is a phase lag. When $t$ gets very large, this pure harmonic oscillation is dominant. The other part of the solution, arising from the general solution of the reduced equation, dies away as $t$ increases. It is called the *transient*, while the solution (3) is called the *steady-state* solution.

## EXERCISES

**1.** Assume that the simple harmonic motion in question is governed by an equation of the type $x'' + \omega^2 x = 0$. Answer the questions, using the given data.

(a) The amplitude is 1, and $x' = 10$ when $x = 0$. Find the period.

(b) The period is $6\pi$, and $x' = 5\sqrt{3}/3$ when $x = 5$. Find the amplitude.

(c) The amplitude is 4. When $x = 2$, $x' = 12\sqrt{3}$. What is the period?

**2.** The vertical oscillations of a ship follow the law $\dfrac{d^2x}{dt^2} + \dfrac{g}{h}x = 0$, where $h$ is the average depth of immersion of the ship. Find the period of vertical oscillation of a ship that draws 10 feet of water.

**3.** Find the steady state oscillation of the system governed by the equation

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + a^2x = E \sin pt,$$

barring a certain exceptional circumstance.

**4.** Find values of $C$ and $D$ so that $x = Ct \cos at + Dt \sin at$ will be a solution of $d^2x/dt^2 + a^2x = E \cos at$. The extra factor $t$ in the solution is needed because the frequency of the impressed force is the same as that of the free simple harmonic oscillations of the system. Observe that the amplitude of the forced vibrations increases with time. This is the phenomenon of *resonance.*

5. Find a particular solution of the equation

$$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + a^2x = Ee^{-\lambda t}\cos pt,$$

assuming $p^2 \neq a^2 - \lambda^2$.

6. Express the solution of the following problem in the form

$$x = Ce^{-\lambda t}\cos(nt - \alpha),$$

finding the values of $C, \lambda, n, \alpha$.

$$\frac{d^2x}{dt^2} + 4\frac{dx}{dt} + 13x = 0;$$

$x = 3\sqrt{3}$ and $\dfrac{dx}{dt} = -3(3 + 2\sqrt{3})$ when $t = 0$.

7. For an electric circuit containing a battery, a resistor, a capacitor, and an inductor in parallel, the current $y$ in one branch of the circuit satisfies the differential equation

$$RLC\frac{d^2y}{dt^2} + L\frac{dy}{dt} + Ry = E,$$

where all the capital letters are constants.
(a) Find the transient and the steady-state solutions in the oscillatory case.
(b) What is the solution in the case of equal roots of the auxiliary equation (the *critically damped* case)?

8. For an electric circuit containing a resistor $R$, an inductor $L$, and a capacitor of capacitance $C$, but no electromotive force, the fundamental equation is

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = 0,$$

where $q$ is the charge on the capacitor at time $t$.
(a) Under what conditions will the discharge be oscillatory?
(b) If $q = q_0$ and $dq/dt = 0$ when $t = 0$, show that

$$q = q_0 \sec\alpha\, e^{-Rt/2L}\cos(nt - \alpha),$$

where $\tan\alpha = \dfrac{R}{2nL}$ and $n = \left(\dfrac{1}{LC} - \dfrac{R^2}{4L^2}\right)^{1/2}$.

The angle $\alpha$ is in the first quadrant.
(c) Find the solution, with the same initial conditions as in (b), when the auxiliary equation has equal roots. Draw the graph of $q$ as a function of $t$, assuming $q_0 > 0$.

## *Review Questions for Chapters XIX, XX, XXI*

CONCEPTS AND DEFINITIONS

1. Define a function of two variables.
2. Give the definition of "$f(x, y) \to A$ as $(x, y) \to (a, b)$." Does $(a, b)$ have to be in the domain of $f$?
3. Define continuity for a function of two variables.
4. What is meant by a level curve of a function? By a level surface?
5. Express the partial derivative of $F(x, y, z)$ with respect to $x$ at $(a, b, c)$ explicitly as a limit.
6. What is the meaning of the statement that a plane $M$ is tangent to a surface $S$ at $(a, b, c)$?
7. Define the meaning of differentiability, and the differential itself, for a function of two variables. What kind of function is the differential?
8. What is the chain rule concerning functions of several variables?
9. What is the invariance of appearance property of differentials?
10. Define absolute maximum and relative maximum for a function of two variables, and indicate the distinction between the two notions.
11. Define the directional derivative concept. Define the gradient of a function. Explain the relations between gradients, directional derivatives, and level surfaces.
12. Define a double integral.
13. Explain the concept of density of a lamina. How is it related to double integrals?
14. What is a compound pendulum? When are the oscillations of a pendulum approximately in simple harmonic motion? Why?
15. Define the area of a smooth surface $z = f(x, y)$.
16. Define a triple integral.
17. State Newton's law of gravitation for the attraction of a solid body on a particle.
18. Explain, with notation, what it means for a function $F$ to be a solution of $dy/dx = f(x, y)$ on a certain interval.
19. What does it mean for the variables to be separable in a differential equation of the first order?
20. What is a direction-field for the differential equation $y' = f(x, y)$. What relation is there between the direction-field and an integral curve?
21. Under what conditions on $f$ is the equation $dy/dx = f(x, y)$ homogeneous?

22. What is the general form of a linear differential equation of the second order? What is meant by the homogeneous equation corresponding to a given equation?

23. Define $e^w$ if $w$ is complex.

24. What is meant by saying that the functions $F_1$, $F_2$ are linearly independent on a certain interval?

THEORY

1. Assuming that the surface $z = f(x, y)$ has a tangent plane at $(a, b, c)$ not parallel to the $z$-axis, derive the equation of this plane.

2. State and prove a version of the chain rule concerning functions of several variables. Do you need the concept of differentiability, or merely the concept of a partial derivative?

3. In what important theorem about extreme values is it important for the relative extreme to occur at an interior point of the domain of a function? State and prove the theorem.

4. Outline a procedure for searching for the absolute maximum or minimum value of a function.

5. What conditions on a function $f$ and a set $S$ are sufficient to guarantee the attainment by $f$ of absolute extremes on $S$?

6. In case $z = f(x, y)$ satisfies a relation $F(x, y, z) = 0$, show how to express the partial derivatives of $f$ in terms of the partial derivatives of $F$, granted certain conditions.

7. Explain in heuristic terms the connection between double integrals and iterated integrals, both interpreted as volumes under a surface $z = f(x, y)$.

8. Outline the procedure for deriving the evaluation of a double integral by an iterated integral in polar coordinates. Among other things, account for the introduction of the factor $r$.

9. Do as in the preceding question for the evaluation of a triple integral by an iterated integral in spherical coordinates. What needs to be accounted for this time?

10. What is the principle of the motion of the center of mass of a system of particles? Prove its validity, using Newton's law for particles.

11. What is the principle of angular momentum? Deduce it from Newton's law for particles. What does it become in the particular case of a rigid system of masses in the $xy$-plane, rotating about the $z$-axis?

12. State and prove some version of the parallel axis theorem for moments of inertia. When is the moment of inertia of a system least, if all axes parallel to a given line are considered?

13. Show how to solve the differential equation $dy/dx = f(y/x)$.

14. Derive the general solution of the equation $y' + P(x)y = Q(x)$. How do you solve an equation of Bernoulli type?

15. Explain how certain second-order differential equations can be solved by first-order methods. Be explicit about the types you consider.

16. Explain the structure of the general solution of a linear differential equation of the second order and of the reduced equation corresponding to it.

17. Using the definition of $e^z$ for complex values of $z$, prove that $e^z e^w = e^{z+w}$.

18. If $f(x)$ is a differentiable complex function of the real variable $x$, prove that

$$\frac{d}{dx} e^{f(x)} = e^{f(x)} f'(x).$$

## *Table I. Table of Natural Logarithms (concluded)*

| | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| **5.5** | 1.7047 | 1.7066 | 1.7084 | 1.7102 | 1.7120 | 1.7138 | 1.7156 | 1.7174 | 1.7192 | 1.7210 |
| 5.6 | .7228 | .7246 | .7263 | .7281 | .7299 | .7317 | .7334 | .7352 | .7370 | .7387 |
| 5.7 | .7405 | .7422 | .7440 | .7457 | .7475 | .7492 | .7509 | .7527 | .7544 | .7561 |
| 5.8 | .7579 | .7596 | .7613 | .7630 | .7647 | .7664 | .7681 | .7699 | .7716 | .7733 |
| 5.9 | .7750 | .7766 | .7783 | .7800 | .7817 | .7843 | .7851 | .7867 | .7884 | .7901 |
| **6.0** | .7918 | .7934 | .7951 | .7967 | .7984 | .8001 | .8017 | .8034 | .8050 | .8066 |
| 6.1 | .8083 | .8099 | .8116 | .8132 | .8148 | .8165 | .8181 | .8197 | .8213 | .8229 |
| 6.2 | .8245 | .8262 | .8278 | .8294 | .8310 | .8326 | .8342 | .8358 | .8374 | .8390 |
| 6.3 | .8405 | .8421 | .8437 | .8453 | .8469 | .8485 | .8500 | .8516 | .8532 | .8547 |
| 6.4 | .8563 | .8579 | .8594 | .8610 | .8625 | .8641 | .8656 | .8672 | .8687 | .8703 |
| **6.5** | .8718 | .8733 | .8749 | .8764 | .8779 | .8795 | .8810 | .8825 | .8840 | .8856 |
| 6.6 | .8871 | .8886 | .8901 | .8916 | .8931 | .8946 | .8961 | .8976 | .8991 | .9006 |
| 6.7 | .9021 | .9036 | .9051 | .9066 | .9081 | .9095 | .9110 | .9125 | .9140 | .9155 |
| 6.8 | .9169 | .9184 | .9199 | .9213 | .9228 | .9242 | .9257 | .9272 | .9286 | .9301 |
| 6.9 | .9315 | .9330 | .9344 | .9359 | .9373 | .9387 | .9402 | .9416 | .9430 | .9445 |
| **7.0** | .9459 | .9473 | .9488 | .9502 | .9516 | .9530 | .9544 | .9559 | .9573 | .9587 |
| 7.1 | .9601 | .9615 | .9629 | .9643 | .9657 | .9671 | .9685 | .9699 | .9713 | .9727 |
| 7.2 | .9741 | .9755 | .9769 | .9782 | .9796 | .9810 | .9824 | .9838 | .9851 | 1.9865 |
| 7.3 | 1.9879 | 1.9892 | 1.9906 | 1.9920 | 1.9933 | 1.9947 | 1.9961 | 1.9974 | 1.9988 | 2.0001 |
| 7.4 | 2.0015 | 2.0028 | 2.0042 | 2.0055 | 2.0069 | 2.0082 | 2.0096 | 2.0109 | 2.0122 | .0136 |
| **7.5** | .0149 | .0162 | .0176 | .0189 | .0202 | .0215 | .0229 | .0242 | .0255 | .0268 |
| 7.6 | .0281 | .0295 | .0308 | .0321 | .0334 | .0347 | .0360 | .0373 | .0386 | .0399 |
| 7.7 | .0412 | .0425 | .0438 | .0451 | .0464 | .0477 | .0490 | .0503 | .0516 | .0528 |
| 7.8 | .0541 | .0554 | .0567 | .0580 | .0592 | .0605 | .0618 | .0631 | .0643 | .0656 |
| 7.9 | .0669 | .0681 | .0694 | .0707 | .0719 | .0732 | .0744 | .0757 | .0769 | .0782 |
| **8.0** | .0794 | .0807 | .0819 | .0832 | .0844 | .0857 | .0869 | .0882 | .0894 | .0906 |
| 8.1 | .0919 | .0931 | .0943 | .0956 | .0968 | .0980 | .0992 | .1005 | .1017 | .1029 |
| 8.2 | .1041 | .1054 | .1066 | .1078 | .1090 | .1102 | .1114 | .1126 | .1138 | .1150 |
| 8.3 | .1163 | .1175 | .1187 | .1199 | .1211 | .1223 | .1235 | .1247 | .1258 | .1270 |
| 8.4 | .1282 | .1294 | .1306 | .1318 | .1330 | .1342 | .1353 | .1365 | .1377 | .1389 |
| **8.5** | .1401 | .1412 | .1424 | .1436 | .1448 | .1459 | .1471 | .1483 | .1494 | .1506 |
| 8.6 | .1518 | .1529 | .1541 | .1552 | .1564 | .1576 | .1587 | .1599 | .1610 | .1622 |
| 8.7 | .1633 | .1645 | .1656 | .1668 | .1679 | .1691 | .1702 | .1713 | .1725 | .1736 |
| 8.8 | .1748 | .1759 | .1770 | .1782 | .1793 | .1804 | .1815 | .1827 | .1838 | .1849 |
| 8.9 | .1861 | .1872 | .1883 | .1894 | .1905 | .1917 | .1928 | .1939 | .1950 | .1961 |
| **9.0** | .1972 | .1983 | .1994 | .2006 | .2017 | .2028 | .2039 | .2050 | .2061 | .2072 |
| 9.1 | .2083 | .2094 | .2105 | .2116 | .2127 | .2138 | .2148 | .2159 | .2170 | .2181 |
| 9.2 | .2192 | .2203 | .2214 | .2225 | .2235 | .2246 | .2257 | .2268 | .2279 | .2289 |
| 9.3 | .2300 | .2311 | .2322 | .2332 | .2343 | .2354 | .2364 | .2375 | .2386 | .2396 |
| 9.4 | .2407 | .2418 | .2428 | .2439 | .2450 | .2460 | .2471 | .2481 | .2492 | .2502 |
| **9.5** | .2513 | .2523 | .2534 | .2544 | .2555 | .2565 | .2576 | .2586 | .2597 | .2607 |
| 9.6 | .2618 | .2628 | .2638 | .2649 | .2659 | .2670 | .2680 | .2690 | .2701 | .2711 |
| 9.7 | .2721 | .2732 | .2742 | .2752 | .2762 | .2773 | .2783 | .2793 | .2803 | .2814 |
| 9.8 | .2824 | .2834 | .2844 | .2854 | .2865 | .2875 | .2885 | .2895 | .2905 | .2915 |
| 9.9 | 2.2925 | 2.2935 | 2.2946 | 2.2956 | 2.2966 | 2.2976 | 2.2986 | 2.2996 | 2.3006 | 2.3016 |
| 10.0 | 2.3026 | | | | | | | | | |

## *Table II. Exponential and Hyperbolic Functions*

| $x$ | $e^x$ | $e^{-x}$ | sinh $x$ | cosh $x$ | tanh $x$ | coth $x$ | $\sinh^{-1} x$ | $\cosh^{-1} x$ | $\tanh^{-1} x$ | $\coth^{-1} x$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.00 | 1.000 | 1.000 | 0.000 | 1.000 | 0.000 | ∞ | 0.000 | ..... | 0.000 | |
| 0.10 | 1.105 | 0.905 | 0.100 | 1.005 | 0.100 | 10.033 | 0.100 | ..... | 0.100 | |
| 0.20 | 1.221 | 0.819 | 0.201 | 1.020 | 0.197 | 5.066 | 0.199 | ..... | 0.203 | |
| 0.30 | 1.350 | 0.741 | 0.305 | 1.045 | 0.291 | 3.433 | 0.296 | ..... | 0.309 | |
| 0.40 | 1.492 | 0.670 | 0.411 | 1.081 | 0.380 | 2.632 | 0.390 | ..... | 0.424 | |
| 0.50 | 1.649 | 0.607 | 0.521 | 1.128 | 0.462 | 2.164 | 0.481 | ..... | 0.549 | |
| 0.60 | 1.822 | 0.549 | 0.637 | 1.185 | 0.537 | 1.862 | 0.569 | ..... | 0.693 | |
| 0.70 | 2.014 | 0.497 | 0.759 | 1.255 | 0.604 | 1.655 | 0.653 | ..... | 0.867 | |
| 0.80 | 2.226 | 0.449 | 0.888 | 1.337 | 0.664 | 1.506 | 0.733 | ..... | 1.099 | |
| 0.90 | 2.460 | 0.407 | 1.027 | 1.433 | 0.716 | 1.396 | 0.809 | ..... | 1.472 | |
| **1.00** | **2.718** | **0.368** | **1.175** | **1.543** | **0.762** | **1.313** | **0.881** | **0.000** | ∞ | ∞ |
| 1.10 | 3.004 | 0.333 | 1.336 | 1.669 | 0.800 | 1.249 | 0.950 | 0.444 | ..... | 1.522 |
| 1.20 | 3.320 | 0.301 | 1.509 | 1.811 | 0.834 | 1.200 | 1.016 | 0.622 | ..... | 1.199 |
| 1.30 | 3.669 | 0.273 | 1.698 | 1.971 | 0.862 | 1.160 | 1.079 | 0.756 | ..... | 1.018 |
| 1.40 | 4.055 | 0.247 | 1.904 | 2.151 | 0.885 | 1.129 | 1.138 | 0.867 | ..... | 0.896 |
| 1.50 | 4.482 | 0.223 | 2.129 | 2.352 | 0.905 | 1.105 | 1.195 | 0.962 | ..... | 0.805 |
| 1.60 | 4.953 | 0.202 | 2.376 | 2.577 | 0.922 | 1.085 | 1.249 | 1.047 | ..... | 0.733 |
| 1.70 | 5.474 | 0.183 | 2.646 | 2.828 | 0.935 | 1.069 | 1.301 | 1.123 | ..... | 0.675 |
| 1.80 | 6.050 | 0.165 | 2.942 | 3.107 | 0.947 | 1.056 | 1.350 | 1.193 | ..... | 0.626 |
| 1.90 | 6.686 | 0.150 | 3.268 | 3.418 | 0.956 | 1.046 | 1.398 | 1.257 | ..... | 0.585 |
| **2.00** | **7.389** | **0.135** | **3.627** | **3.762** | **0.964** | **1.037** | **1.444** | **1.317** | ..... | **0.549** |
| 2.10 | 8.166 | 0.122 | 4.022 | 4.144 | 0.970 | 1.030 | 1.487 | 1.373 | ..... | 0.518 |
| 2.20 | 9.025 | 0.111 | 4.457 | 4.568 | 0.976 | 1.025 | 1.530 | 1.425 | ..... | 0.490 |
| 2.30 | 9.974 | 0.100 | 4.937 | 5.037 | 0.980 | 1.020 | 1.570 | 1.475 | ..... | 0.466 |
| 2.40 | 11.02 | 0.091 | 5.466 | 5.557 | 0.984 | 1.017 | 1.609 | 1.522 | ..... | 0.444 |
| 2.50 | 12.18 | 0.082 | 6.050 | 6.132 | 0.987 | 1.014 | 1.647 | 1.567 | ..... | 0.424 |
| 2.60 | 13.46 | 0.074 | 6.695 | 6.769 | 0.989 | 1.011 | 1.684 | 1.609 | ..... | 0.405 |
| 2.70 | 14.88 | 0.067 | 7.406 | 7.473 | 0.991 | 1.009 | 1.719 | 1.650 | ..... | 0.389 |
| 2.80 | 16.44 | 0.061 | 8.192 | 8.253 | 0.993 | 1.007 | 1.753 | 1.689 | ..... | 0.374 |
| 2.90 | 18.17 | 0.055 | 9.060 | 9.115 | 0.994 | 1.006 | 1.786 | 1.727 | ..... | 0.360 |
| **3.00** | **20.09** | **0.050** | **10.018** | **10.068** | **0.995** | **1.005** | **1.818** | **1.763** | ..... | **0.347** |
| 3.10 | 22.20 | 0.045 | 11.08 | 11.12 | 0.996 | 1.004 | 1.850 | 1.798 | ..... | 0.335 |
| 3.20 | 24.53 | 0.041 | 12.25 | 12.29 | 0.997 | 1.003 | 1.880 | 1.831 | ..... | 0.323 |
| 3.30 | 27.11 | 0.037 | 13.54 | 13.57 | 0.997 | 1.003 | 1.909 | 1.863 | ..... | 0.313 |
| 3.40 | 29.96 | 0.033 | 14.97 | 15.00 | 0.998 | 1.002 | 1.938 | 1.895 | ..... | 0.303 |
| 3.50 | 33.12 | 0.030 | 16.54 | 16.57 | 0.998 | 1.002 | 1.966 | 1.925 | ..... | 0.294 |
| 4.00 | 54.60 | 0.018 | 27.29 | 27.31 | 0.999 | 1.001 | 2.095 | 2.063 | ..... | 0.255 |
| 4.50 | 90.02 | 0.0111 | 45.00 | 45.01 | 1.000 | 1.000 | 2.209 | 2.185 | ..... | 0.226 |
| **5.00** | **148.4** | **0.0067** | **74.20** | **74.21** | **1.000** | **1.000** | **2.312** | **2.292** | ..... | **0.203** |
| 5.50 | 244.7 | 0.0041 | 122.3 | 122.3 | 1.000 | 1.000 | 2.406 | 2.390 | ..... | 0.184 |
| 6.00 | 403.4 | 0.0025 | 201.7 | 201.7 | 1.000 | 1.000 | 2.492 | 2.478 | ..... | 0.168 |
| $x$ | $e^x$ | $e^{-x}$ | sinh $x$ | cosh $x$ | tanh $x$ | coth $x$ | $\sinh^{-1} x$ | $\cosh^{-1} x$ | $\tanh^{-1} x$ | $\coth^{-1} x$ |

## *Table III. Natural Functions for Angles in Radians*

| x | sin x | tan x | ctn x | cos x | x | sin x | tan x | ctn x | cos x |
|---|---|---|---|---|---|---|---|---|---|
| **.00** | .00000 | .00000 | None | 1.0000 | **.40** | .38942 | .42279 | 2.3652 | .92106 |
| .01 | .01000 | .01000 | 99.997 | .99995 | .41 | .39861 | .43463 | 2.3008 | .91712 |
| .02 | .02000 | .02000 | 49.993 | .99980 | .42 | .40776 | .44657 | 2.2393 | .91309 |
| .03 | .03000 | .03001 | 33.323 | .99955 | .43 | .41687 | .45862 | 2.1804 | .90897 |
| .04 | .03999 | .04002 | 24.987 | .99920 | .44 | .42594 | .47078 | 2.1241 | .90475 |
| **.05** | .04998 | .05004 | 19.983 | .99875 | **.45** | .43497 | .48306 | 2.0702 | .90045 |
| .06 | .05996 | .06007 | 16.647 | .99820 | .46 | .44395 | .49545 | 2.0184 | .89605 |
| .07 | .06994 | .07011 | 14.262 | .99755 | .47 | .45289 | .50797 | 1.9686 | .89157 |
| .08 | .07991 | .08017 | 12.473 | .99680 | .48 | .46178 | .52061 | 1.9208 | .88699 |
| .09 | .08998 | .09024 | 11.081 | .99595 | .49 | .47063 | .53339 | 1.8748 | .88233 |
| **.10** | .09983 | .10033 | 9.9666 | .99500 | **.50** | .47943 | .54630 | 1.8305 | .87758 |
| .11 | .10978 | .11045 | 9.0542 | .99396 | .51 | .48818 | .55936 | 1.7878 | .87274 |
| .12 | .11971 | .12058 | 8.2933 | .99281 | .52 | .49688 | .57256 | 1.7465 | .86782 |
| .13 | .12963 | .13074 | 7.6489 | .99156 | .53 | .50553 | .58592 | 1.7067 | .86281 |
| .14 | .13954 | .14092 | 7.0961 | .99022 | .54 | .51414 | .59943 | 1.6683 | .85771 |
| **.15** | .14944 | .15114 | 6.6166 | .98877 | **.55** | .52269 | .61311 | 1.6310 | .85252 |
| .16 | .15932 | .16138 | 6.1966 | .98723 | .56 | .53119 | .62695 | 1.5950 | .84726 |
| .17 | .16918 | .17166 | 5.8256 | .98558 | .57 | .53963 | .64097 | 1.5601 | .84190 |
| .18 | .17903 | .18197 | 5.4954 | .98384 | .58 | .54802 | .64517 | 1.5263 | .83646 |
| .19 | .18886 | .19232 | 5.1997 | .98200 | .59 | .55636 | .66956 | 1.4935 | .83094 |
| **.20** | .19867 | .20271 | 4.9332 | .98007 | **.60** | .56464 | .68414 | 1.4617 | .82534 |
| .21 | .20846 | .21314 | 4.6917 | .97803 | .61 | .57287 | .69892 | 1.4308 | .81965 |
| .22 | .21823 | .22362 | 4.4719 | .97590 | .62 | .58104 | .71391 | 1.4007 | .81388 |
| .23 | .22798 | .23414 | 4.2709 | .97367 | .63 | .58914 | .72911 | 1.3715 | .80803 |
| .24 | .23770 | .24472 | 4.0864 | .97134 | .64 | .59720 | .74454 | 1.3431 | .80210 |
| **.25** | .24740 | .25534 | 3.9163 | .96891 | **.65** | .60519 | .76020 | 1.3154 | .79608 |
| .26 | .25708 | .26602 | 3.7591 | .96639 | .66 | .61312 | .77610 | 1.2885 | .78999 |
| .27 | .26673 | .27676 | 3.6133 | .96377 | .67 | .62099 | .79225 | 1.2622 | .78382 |
| .28 | .27636 | .28755 | 3.4776 | .96106 | .68 | .62879 | .80866 | 1.2366 | .77757 |
| .29 | .28595 | .29841 | 3.3511 | .95824 | .69 | .63654 | .82534 | 1.2116 | .77125 |
| **.30** | .29552 | .30934 | 3.2327 | .95534 | **.70** | .64422 | .84229 | 1.1872 | .76484 |
| .31 | .30506 | .32033 | 3.1218 | .95233 | .71 | .65183 | .85953 | 1.1634 | .75836 |
| .32 | .31457 | .33139 | 3.0176 | .94924 | .72 | .65938 | .87707 | 1.1402 | .75181 |
| .33 | .32404 | .34252 | 2.9195 | .94604 | .73 | .66687 | .89492 | 1.1174 | .74517 |
| .34 | .33349 | .35374 | 2.8270 | .94275 | .74 | .67429 | .91309 | 1.0952 | .73847 |
| **.35** | .34290 | .36503 | 2.7395 | .93937 | **.75** | .68164 | .93160 | 1.0734 | .73169 |
| .36 | .35227 | .37640 | 2.6567 | .93590 | .76 | .68892 | .95045 | 1.0521 | .72484 |
| .37 | .36162 | .38786 | 2.5782 | .93233 | .77 | .69614 | 6.96967 | 1.0313 | .71791 |
| .38 | .37092 | .39941 | 2.5037 | .92866 | .78 | .70328 | .98926 | 1.0109 | .71091 |
| .39 | .38019 | .41105 | 2.4328 | .92491 | .79 | .71035 | 1.0092 | .99084 | .70385 |

## *Table III. Natural Functions for Angles in Radians (concluded)*

| $x$ | sin $x$ | tan $x$ | ctn $x$ | cos $x$ | $x$ | sin $x$ | tan $x$ | ctn $x$ | cos $x$ |
|---|---|---|---|---|---|---|---|---|---|
| **.80** | .71736 | 1.0296 | .97121 | .69671 | **1.20** | .93204 | 2.5722 | .38878 | .36236 |
| .81 | .72429 | 1.0505 | .95197 | .68950 | 1.21 | .93562 | 2.6503 | .37731 | .35302 |
| .82 | .73115 | 1.0717 | .93309 | .68222 | 1.22 | .93910 | 2.7328 | .36593 | .34365 |
| .83 | .73793 | 1.0934 | .91455 | .67488 | 1.23 | .94249 | 2.8198 | .35463 | .33424 |
| .84 | .74464 | 1.1156 | .89635 | .66746 | 1.24 | .94578 | 2.9119 | .34341 | .32480 |
| **.85** | .75128 | 1.1383 | .87848 | .65998 | **1.25** | .94898 | 3.0096 | .33227 | .31532 |
| .86 | .75784 | 1.1616 | .86091 | .65244 | 1.26 | .95209 | 3.1133 | .32121 | .30582 |
| .87 | .76433 | 1.1853 | .84365 | .64483 | 1.27 | .95510 | 3.2236 | .31021 | .29628 |
| .88 | .77074 | 1.2097 | .82668 | .63715 | 1.28 | .95802 | 3.3413 | .29928 | .28672 |
| .89 | .77707 | 1.2346 | .80998 | .62941 | 1.29 | .96084 | 3.4672 | .28842 | .27712 |
| **.90** | .78333 | 1.2602 | .79355 | .62161 | **1.30** | .96356 | 3.6021 | .27762 | .26750 |
| .91 | .78950 | 1.2864 | .77738 | .61375 | 1.31 | .96618 | 3.7471 | .26687 | .25785 |
| .92 | .79560 | 1.3133 | .76146 | .60582 | 1.32 | .96872 | 3.9033 | .25619 | .24818 |
| .93 | .80162 | 1.3409 | .74578 | .59783 | 1.33 | .97115 | 4.0723 | .24556 | .23848 |
| .94 | .80756 | 1.3692 | .73034 | .58979 | 1.34 | .97348 | 4.2556 | .23498 | .22875 |
| **.95** | .81342 | 1.3984 | .71511 | .58168 | **1.35** | .97572 | 4.4552 | .22446 | .21901 |
| .96 | .81919 | 1.4284 | .70010 | .57352 | 1.36 | .97786 | 4.6734 | .21398 | .20924 |
| .97 | .82489 | 1.4592 | .68531 | .56530 | 1.37 | .97991 | 4.9131 | .20354 | .19945 |
| .98 | .83050 | 1.4910 | .67071 | .55702 | 1.38 | .98185 | 5.1774 | .19315 | .18964 |
| .99 | .83603 | 1.5237 | .65631 | .54869 | 1.39 | .98370 | 5.4707 | .18279 | .17981 |
| **1.00** | .84147 | 1.5574 | .64209 | .54030 | **1.40** | .98545 | 5.7979 | .17248 | .16997 |
| 1.01 | .84633 | 1.5922 | .62806 | .53186 | 1.41 | .98710 | 6.1654 | .16220 | .16010 |
| 1.02 | .85211 | 1.6281 | .61420 | .52337 | 1.42 | .98865 | 6.5811 | .15195 | .15023 |
| 1.03 | .85730 | 1.6652 | .60051 | .51482 | 1.43 | .99010 | 7.0555 | .14173 | .14033 |
| 1.04 | .86240 | 1.7036 | .58699 | .50622 | 1.44 | .99146 | 7.6018 | .13155 | .13042 |
| **1.05** | .86742 | 1.7433 | .57362 | .49757 | **1.45** | .99271 | 8.2381 | .12139 | .12050 |
| 1.06 | .87236 | 1.7844 | .56040 | .48887 | 1.46 | .99387 | 8.9886 | .11125 | .11057 |
| 1.07 | .87720 | 1.8270 | .54734 | .48012 | 1.47 | .99492 | 9.8874 | .10114 | .10063 |
| 1.08 | .88196 | 1.8712 | .53441 | .47133 | 1.48 | .99588 | 10.983 | .09105 | .09067 |
| 1.09 | .88663 | 1.9171 | .52162 | .46249 | 1.49 | .99674 | 12.350 | .08097 | .08071 |
| **1.10** | .89121 | 1.9648 | .50897 | .45360 | **1.50** | .99749 | 14.101 | .07091 | .07074 |
| 1.11 | .89570 | 2.0143 | .49644 | .44466 | 1.51 | .99815 | 16.428 | .06087 | .06076 |
| 1.12 | .90010 | 2.0660 | .48404 | .43568 | 1.52 | .99871 | 19.670 | .05084 | .05077 |
| 1.13 | .90441 | 2.1198 | .47175 | .42666 | 1.53 | .99917 | 24.498 | .04082 | .04079 |
| 1.14 | .90863 | 2.1759 | .45959 | .41759 | 1.54 | .99953 | 32.461 | .03081 | .03079 |
| **1.15** | .91276 | 2.2345 | .44753 | .40849 | **1.55** | .99978 | 48.078 | .02080 | .02079 |
| 1.16 | .91680 | 2.2958 | .43558 | .39934 | 1.56 | .99994 | 92.621 | .01080 | .01080 |
| 1.17 | .92075 | 2.3600 | .42373 | .39015 | 1.57 | 1.0000 | 1255.8 | .00080 | .00080 |
| 1.18 | .92461 | 2.4273 | .41199 | .38092 | 1.58 | .99996 | −108.65 | −.00920 | −.00920 |
| 1.19 | .92837 | 2.4979 | .40034 | .37166 | 1.59 | .99982 | −52.067 | −.01921 | −.01920 |
| | | | | | **1.60** | .99957 | −34.233 | −.02921 | −.02920 |

## *Table IV. Values of Trigonometric Functions*

| Degrees | Radians | Sin | Csc | Tan | Cot | Sec | Cos | | |
|---|---|---|---|---|---|---|---|---|---|
| **0° 0′** | .0000 | .0000 | — | .0000 | — | 1.000 | 1.0000 | 1.5708 | **90° 0′** |
| 10′ | 029 | 029 | 343.8 | 029 | 343.8 | 000 | 000 | 679 | 50′ |
| 20′ | 058 | 058 | 171.9 | 058 | 171.9 | 000 | 000 | 650 | 40′ |
| **30′** | .0087 | .0087 | 114.6 | .0087 | 114.6 | 1.000 | 1.0000 | 1.5621 | **30′** |
| 40′ | 116 | 116 | 85.95 | 116 | 85.94 | 000 | 9999 | 592 | 20′ |
| 50′ | 145 | 145 | 68.76 | 145 | 68.75 | 000 | 999 | 563 | 10′ |
| **1° 0′** | .0175 | .0175 | 57.30 | .0175 | 57.29 | 1.000 | .9998 | 1.5533 | **89° 0′** |
| 10′ | 204 | 204 | 49.11 | 204 | 49.10 | 000 | 998 | 504 | 50′ |
| 20′ | 233 | 233 | 42.98 | 233 | 42.96 | 000 | 997 | 475 | 40′ |
| **30′** | .0262 | .0262 | 38.20 | .0262 | 38.19 | 1.000 | .9997 | 1.5446 | **30′** |
| 40′ | 291 | 291 | 34.38 | 291 | 34.37 | 000 | 996 | 417 | 20′ |
| 50′ | 320 | 320 | 31.26 | 320 | 31.24 | 001 | 995 | 388 | 10′ |
| **2° 0′** | .0349 | .0349 | 28.65 | .0349 | 28.64 | 1.001 | .9994 | 1.5359 | **88° 0′** |
| 10′ | 378 | 378 | 26.45 | 378 | 26.43 | 001 | 993 | 330 | 50′ |
| 20′ | 407 | 407 | 24.56 | 407 | 24.54 | 001 | 992 | 301 | 40′ |
| **30′** | .0436 | .0436 | 22.93 | .0437 | 22.90 | 1.001 | .9990 | 1.5272 | **30′** |
| 40′ | 465 | 465 | 21.49 | 466 | 21.47 | 001 | 989 | 243 | 20′ |
| 50′ | 495 | 494 | 20.23 | 495 | 20.21 | 001 | 988 | 213 | 10′ |
| **3° 0′** | .0524 | .0523 | 19.11 | .0524 | 19.08 | 1.001 | .9986 | 1.5184 | **87° 0′** |
| 10′ | 553 | 552 | 18.10 | 553 | 18.07 | 002 | 985 | 155 | 50′ |
| 20′ | 582 | 581 | 17.20 | 582 | 17.17 | 002 | 983 | 126 | 40′ |
| **30′** | .0611 | .0610 | 16.38 | .0612 | 16.35 | 1.002 | .9981 | 1.5097 | **30′** |
| 40′ | 640 | 640 | 15.64 | 641 | 15.60 | 002 | 980 | 068 | 20′ |
| 50′ | 669 | 669 | 14.96 | 670 | 14.92 | 002 | 978 | 039 | 10′ |
| **4° 0′** | .0698 | .0698 | 14.34 | .0699 | 14.30 | 1.002 | .9976 | 1.5010 | **86° 0′** |
| 10′ | 727 | 727 | 13.76 | 729 | 13.73 | 003 | 974 | 981 | 50′ |
| 20′ | 756 | 756 | 13.23 | 758 | 13.20 | 003 | 971 | 952 | 40′ |
| **30′** | .0785 | .0785 | 12.75 | .0787 | 12.71 | 1.003 | .9969 | 1.4923 | **30′** |
| 40′ | 814 | 814 | 12.29 | 816 | 12.25 | 003 | 967 | 893 | 20′ |
| 50′ | 844 | 843 | 11.87 | 846 | 11.83 | 004 | 964 | 864 | 10′ |
| **5° 0′** | .0873 | .0872 | 11.47 | .0875 | 11.43 | 1.004 | .9962 | 1.4835 | **85° 0′** |
| 10′ | 902 | 901 | 11.10 | 904 | 11.06 | 004 | 959 | 806 | 50′ |
| 20′ | 931 | 929 | 10.76 | 934 | 10.71 | 004 | 957 | 777 | 40′ |
| **30′** | .0960 | .0958 | 10.43 | .0963 | 10.39 | 1.005 | .9954 | 1.4748 | **30′** |
| 40′ | 989 | 987 | 10.13 | 992 | 10.08 | 005 | 951 | 719 | 20′ |
| 50′ | .1018 | .1016 | 9.839 | .1022 | 9.788 | 005 | 948 | 690 | 10′ |
| **6° 0′** | .1047 | .1045 | 9.567 | .1051 | 9.514 | 1.006 | .9945 | 1.4661 | **84° 0′** |
| 10′ | 076 | 074 | 9.309 | 080 | 9.255 | 006 | 942 | 632 | 50′ |
| 20′ | 105 | 103 | 9.065 | 110 | 9.010 | 006 | 939 | 603 | 40′ |
| **30′** | .1134 | .1132 | 8.834 | .1139 | 8.777 | 1.006 | .9936 | 1.4573 | **30′** |
| 40′ | 164 | 161 | 8.614 | 169 | 8.556 | 007 | 932 | 544 | 20′ |
| 50′ | 193 | 190 | 8.405 | 198 | 8.345 | 007 | 929 | 515 | 10′ |
| **7° 0′** | .1222 | .1219 | 8.206 | .1228 | 8.144 | 1.008 | .9925 | 1.4486 | **83° 0′** |
| 10′ | 251 | 248 | 8.016 | 257 | 7.953 | 008 | 922 | 457 | 50′ |
| 20′ | 280 | 276 | 7.834 | 287 | 7.770 | 008 | 918 | 428 | 40′ |
| **30′** | 1309 | .1305 | 7.661 | .1317 | 7.596 | 1.009 | .9914 | 1.4399 | **30′** |
| 40′ | 338 | 334 | 7.496 | 346 | 7.429 | 009 | 911 | 370 | 20′ |
| 50′ | 367 | 363 | 7.337 | 376 | 7.269 | 009 | 907 | 341 | 10′ |
| **8° 0′** | .1396 | .1392 | 7.185 | .1405 | 7.115 | 1.010 | .9903 | 1.4312 | **82° 0′** |
| 10′ | 425 | 421 | 7.040 | 435 | 6.968 | 010 | 899 | 283 | 50′ |
| 20′ | 454 | 449 | 6.900 | 465 | 6.827 | 011 | 894 | 254 | 40′ |
| **30′** | .1484 | .1478 | 6.765 | .1495 | 6.691 | 1.011 | .9890 | 1.4224 | **30′** |
| 40′ | 513 | 507 | 6.636 | 524 | 6.561 | 012 | 886 | 195 | 20′ |
| 50′ | 542 | 536 | 6.512 | 554 | 6.435 | 012 | 881 | 166 | 10′ |
| **9° 0′** | .1571 | .1564 | 6.392 | .1584 | 6.314 | 1.012 | .9877 | 1.4137 | **81° 0′** |
| | | **Cos** | **Sec** | **Cot** | **Tan** | **Csc** | **Sin** | **Radians** | **Degrees** |

## *Table IV. Values of Trigonometric Functions (continued)*

| Degrees | Radians | Sin | Csc | Tan | Cot | Sec | Cos | | |
|---|---|---|---|---|---|---|---|---|---|
| **9° 0′** | .1571 | .1564 | 6.392 | .1584 | 6.314 | 1.012 | .9877 | 1.4137 | **81° 0′** |
| 10′ | 600 | 593 | 277 | 614 | 197 | 013 | 872 | 108 | 50′ |
| 20′ | 629 | 622 | 166 | 644 | 084 | 013 | 868 | 079 | 40′ |
| **30′** | .1658 | .1650 | 6.059 | .1673 | 5.976] | 1.014 | .9863 | 1.4050 | **30′** |
| 40′ | 687 | 679 | 5.955 | 703 | 871 | 014 | 858 | 1.4021 | 20′ |
| 50′ | 716 | 708 | 855 | 733 | 769 | 015 | 853 | 992 | 10′ |
| **10° 0′** | .1745 | .1736 | 5.759 | .1763 | 5.671 | 1.015 | .9848 | 1.3963 | **80° 0′** |
| 10′ | 774 | 765 | 665 | 793 | 576 | 016 | 843 | 934 | 50′ |
| 20′ | 804 | 794 | 575 | 823 | 485 | 016 | 838 | 904 | 40′ |
| **30′** | .1833 | .1822 | 5.487 | .1853 | 5.396 | 1.017 | .9833 | 1.3875 | **30′** |
| 40′ | 862 | 851 | 403 | 883 | 309 | 018 | 827 | 846 | 20′ |
| 50′ | 891 | 880 | 320 | 914 | 226 | 018 | 822 | 817 | 10′ |
| **11° 0′** | .1920 | .1908 | 5.241 | .1944 | 5.145 | 1.019 | .9816 | 1.3788 | **79° 0′** |
| 10′ | 949 | 937 | 164 | 974 | 066 | 019 | 811 | 759 | 50′ |
| 20′ | 978 | 965 | 089 | .2004 | 4.989 | 020 | 805 | 730 | 40′ |
| **30′** | .2007 | .1994 | 5.016 | .2035 | 4.915 | 1.020 | .9799 | 1.3701 | **30′** |
| 40′ | 036 | .2022 | 4.945 | 065 | 843 | 021 | 793 | 672 | 20′ |
| 50′ | 065 | 051 | 876 | 095 | 773 | 022 | 787 | 643 | 10′ |
| **12° 0′** | .2094 | .2079 | 4.810 | .2126 | 4.705 | 1.022 | .9781 | 1.3614 | **78° 0′** |
| 10′ | 123 | 108 | 745 | 156 | 638 | 023 | 775 | 584 | 50′ |
| 20′ | 153 | 136 | 682 | 186 | 574 | 024 | 769 | 555 | 40′ |
| **30′** | .2182 | .2164 | 4.620 | .2217 | 4.511 | 1.024 | .9763 | 1.3526 | **30′** |
| 40′ | 211 | 193 | 560 | 247 | 449 | 025 | 757 | 497 | 20′ |
| 50′ | 240 | 221 | 502 | 278 | 390 | 026 | 750 | 468 | 10′ |
| **13° 0′** | .2269 | .2250 | 4.445 | .2309 | 4.331 | 1.026 | .9744 | 1.3439 | **77° 0′** |
| 10′ | 298 | 278 | 390 | 339 | 275 | 027 | 737 | 410 | 50′ |
| 20′ | 327 | 306 | 336 | 370 | 219 | 028 | 730 | 381 | 40′ |
| **30′** | .2356 | .2334 | 4.284 | .2401 | 4.165 | 1.028 | .9724 | 1.3352 | **30′** |
| 40′ | 385 | 363 | 232 | 432 | 113 | 029 | 717 | 323 | 20′ |
| 50′ | 414 | 391 | 182 | 462 | 061 | 030 | 710 | 294 | 10′ |
| **14° 0′** | .2443 | .2419 | 4.134 | .2493 | 4.011 | 1.031 | .9703 | 1.3265 | **76° 0′** |
| 10′ | 473 | 447 | 086 | 524 | 3.962 | 031 | 696 | 235 | 50′ |
| 20′ | 502 | 476 | 039 | 555 | 914 | 032 | 689 | 206 | 40′ |
| **30′** | .2531 | .2504 | 3.994 | .2586 | 3.867 | 1.033 | .9681 | 1.3177 | **30′** |
| 40′ | 560 | 532 | 950 | 617 | 821 | 034 | 674 | 148 | 20′ |
| 50′ | 589 | 560 | 906 | 648 | 776 | 034 | 667 | 119 | 10′ |
| **15° 0′** | .2618 | .2588 | 3.864 | .2679 | 3.732 | 1.035 | .9659 | 1.3090 | **75° 0′** |
| 10′ | 647 | 616 | 822 | 711 | 689 | 036 | 652 | 061 | 50′ |
| 20′ | 676 | 644 | 782 | 742 | 647 | 037 | 644 | 032 | 40′ |
| **30′** | .2705 | .2672 | 3.742 | .2773 | 3.606 | 1.038 | .9636 | 1.3003 | **30′** |
| 40′ | 734 | 700 | 703 | 805 | 566 | 039 | 628 | 974 | 20′ |
| 50′ | 763 | 728 | 665 | 836 | 526 | 039 | 621 | 945 | 10′ |
| **16° 0′** | .2793 | .2756 | 3.628 | .2867 | 3.487 | 1.040 | .9613 | 1.2915 | **74° 0′** |
| 10′ | 822 | 784 | 592 | 899 | 450 | 041 | 605 | 886 | 50′ |
| 20′ | 851 | 812 | 556 | 931 | 412 | 042 | 596 | 857 | 40′ |
| **30′** | .2880 | .2840 | 3.521 | .2962 | 3.376 | 1.043 | .9588 | 1.2828 | **30′** |
| 40′ | 909 | 868 | 487 | 994 | 340 | 044 | 580 | 799 | 20′ |
| 50′ | 938 | 896 | 453 | .3026 | 305 | 045 | 572 | 770 | 10′ |
| **17° 0′** | .2967 | .2924 | 3.420 | .3057 | 3.271 | 1.046 | .9563 | 1.2741 | **73° 0′** |
| 10′ | 996 | 952 | 388 | 089 | 237 | 047 | 555 | 712 | 50′ |
| 20′ | .3025 | 979 | 357 | 121 | 204 | 048 | 546 | 683 | 40′ |
| **30′** | .3054 | .3007 | 3.326 | .3153 | 3.172 | 1.048 | .9537 | 1.2654 | **30′** |
| 40′ | 083 | 035 | 295 | 185 | 140 | 049 | 528 | 625 | 20′ |
| 50′ | 113 | 062 | 265 | 217 | 108 | 050 | 520 | 595 | 10′ |
| **18° 0′** | .3142 | .3090 | 3.236 | .3249 | 3.078 | 1.051 | .9511 | 1.2566 | **72° 0′** |
| | | **Cos** | **Sec** | **Cot** | **Tan** | **Csc** | **Sin** | **Radians** | **Degrees** |

## *Table IV. Values of Trigonometric Functions (continued)*

| Degrees | Radians | Sin | Csc | Tan | Cot | Sec | Cos | | |
|---|---|---|---|---|---|---|---|---|---|
| **18° 0′** | .3142 | .3090 | 3.236 | .3249 | 3.078 | 1.051 | .9511 | 1.2566 | **72° 0′** |
| 10′ | 171 | 118 | 207 | 281 | 047 | 052 | 502 | 537 | 50′ |
| 20′ | 200 | 145 | 179 | 314 | 018 | 053 | 492 | 508 | 40′ |
| **30′** | .3229 | .3173 | 3.152 | .3346 | 2.989 | 1.054 | .9483 | 1.2479 | **30′** |
| 40′ | 258 | 201 | 124 | 378 | 960 | 056 | 474 | 450 | 20′ |
| 50′ | 287 | 228 | 098 | 411 | 932 | 057 | 465 | 421 | 10′ |
| **19° 0′** | .3316 | .3256 | 3.072 | .3443 | 2.904 | 1.058 | .9455 | 1.2392 | **71° 0′** |
| 10′ | 345 | 283 | 046 | 476 | 877 | 059 | 446 | 363 | 50′ |
| 20′ | 374 | 311 | 021 | 508 | 850 | 060 | 436 | 334 | 40′ |
| **30′** | .3403 | .3338 | 2.996 | .3541 | 2.824 | 1.061 | .9426 | 1.2305 | **30′** |
| 40′ | 432 | 365 | 971 | 574 | 798 | 062 | 417 | 275 | 20′ |
| 50′ | 462 | 393 | 947 | 607 | 773 | 063 | 407 | 246 | 10′ |
| **20° 0′** | .3491 | .3420 | 2.924 | .3640 | 2.747 | 1.064 | .9397 | 1.2217 | **70° 0′** |
| 10′ | 520 | 448 | 901 | 673 | 723 | 065 | 387 | 188 | 50′ |
| 20′ | 549 | 475 | 878 | 706 | 699 | 066 | 377 | 159 | 40′ |
| **30′** | .3578 | .3502 | 2.855 | .3739 | 2.675 | 1.068 | .9367 | 1.2130 | **30′** |
| 40′ | 607 | 529 | 833 | 772 | 651 | 069 | 356 | 101 | 20′ |
| 50′ | 636 | 557 | 812 | 805 | 628 | 070 | 346 | 072 | 10′ |
| **21° 0′** | .3665 | .3584 | 2.790 | .3839 | 2.605 | 1.071 | .9336 | 1.2043 | **69° 0′** |
| 10′ | 694 | 611 | 769 | 872 | 583 | 072 | 325 | 1.2014 | 50′ |
| 20′ | 723 | 638 | 749 | 906 | 560 | 074 | 315 | 985 | 40′ |
| **30′** | .3752 | .3665 | 2.729 | .3939 | 2.539 | 1.075 | .9304 | 1.1956 | **30′** |
| 40′ | 782 | 692 | 709 | 973 | 517 | 076 | 293 | 926 | 20′ |
| 50′ | 811 | 719 | 689 | .4006 | 496 | 077 | 283 | 897 | 10′ |
| **22° 0′** | .3840 | .3746 | 2.669 | .4040 | 2.475 | 1.079 | .9272 | 1.1868 | **68° 0′** |
| 10′ | 869 | 773 | 650 | 074 | 455 | 080 | 261 | 839 | 50′ |
| 20′ | 898 | 800 | 632 | 108 | 434 | 081 | 250 | 810 | 40′ |
| **30′** | .3927 | .3827 | 2.613 | .4142 | 2.414 | 1.082 | .9239 | 1.1781 | **30′** |
| 40′ | 956 | 854 | 595 | 176 | 394 | 084 | 228 | 752 | 20′ |
| 50′ | 985 | 881 | 577 | 210 | 375 | 085 | 216 | 723 | 10′ |
| **23° 0′** | .4014 | .3907 | 2.559 | .4245 | 2.356 | 1.086 | .9205 | 1.1694 | **67° 0′** |
| 10′ | 043 | 934 | 542 | 279 | 337 | 088 | 194 | 665 | 50′ |
| 20′ | 072 | 961 | 525 | 314 | 318 | 089 | 182 | 636 | 40′ |
| **30′** | .4102 | .3987 | 2.508 | .4348 | 2.300 | 1.090 | .9171 | 1.1606 | **30′** |
| 40′ | 131 | .4014 | 491 | 383 | 282 | 092 | 159 | 577 | 20′ |
| 50′ | 160 | 041 | 475 | 417 | 264 | 093 | 147 | 548 | 10′ |
| **24° 0′** | .4189 | .4067 | 2.459 | .4452 | 2.246 | 1.095 | .9135 | 1.1519 | **66° 0′** |
| 10′ | 218 | 094 | 443 | 487 | 229 | 096 | 124 | 490 | 50′ |
| 20′ | 247 | 120 | 427 | 522 | 211 | 097 | 112 | 461 | 40′ |
| **30′** | .4276 | .4147 | 2.411 | .4557 | 2.194 | 1.099 | .9100 | 1.1432 | **30′** |
| 40′ | 305 | 173 | 396 | 592 | 177 | 100 | 088 | 403 | 20′ |
| 50′ | 334 | 200 | 381 | 628 | 161 | 102 | 075 | 374 | 10′ |
| **25° 0′** | .4363 | .4226 | 2.366 | .4663 | 2.145 | 1.103 | .9063 | 1.1345 | **65° 0′** |
| 10′ | 392 | 253 | 352 | 699 | 128 | 105 | 051 | 316 | 50′ |
| 20′ | 422 | 279 | 337 | 734 | 112 | 106 | 038 | 286 | 40′ |
| **30′** | .4451 | .4305 | 2.323 | .4770 | 2.097 | 1.108 | .9026 | 1.1257 | **30′** |
| 40′ | 480 | 331 | 309 | 806 | 081 | 109 | 013 | 228 | 20′ |
| 50′ | 509 | 358 | 295 | 841 | 066 | 111 | 001 | 199 | 10′ |
| **26° 0′** | .4538 | .4384 | 2.281 | .4877 | 2.050 | 1.113 | .8988 | 1.1170 | **64° 0′** |
| 10′ | 567 | 410 | 268 | 913 | 035 | 114 | 975 | 141 | 50′ |
| 20′ | 596 | 436 | 254 | 950 | 020 | 116 | 962 | 112 | 40′ |
| **30′** | .4625 | .4462 | 2.241 | .4986 | 2.006 | 1.117 | .8949 | 1.1083 | **30′** |
| 40′ | 654 | 488 | 228 | .5022 | 1.991 | 119 | 936 | 054 | 20′ |
| 50′ | 683 | 514 | 215 | 059 | 977 | 121 | 923 | 1.1025 | 10′ |
| **27° 0′** | .4712 | .4540 | 2.203 | .5095 | 1.963 | 1.122 | .8910 | 1.0996 | **63° 0′** |
| | | Cos | Sec | Cot | Tan | Csc | Sin | Radians | Degrees |

## Table IV. Values of Trigonometric Functions (continued)

| Degrees | Radians | Sin | Csc | Tan | Cot | Sec | Cos | | |
|---|---|---|---|---|---|---|---|---|---|
| **27° 0′** | .4712 | .4540 | 2.263 | .5095 | 1.963 | 1.122 | .8910 | 1.0996 | **63° 0′** |
| 10′ | 741 | 566 | 190 | 132 | 949 | 124 | 897 | 966 | 50′ |
| 20′ | 771 | 592 | 178 | 169 | 935 | 126 | 884 | 937 | 40′ |
| **30′** | .4800 | .4617 | 2.166 | .5206 | 1.921 | 1.127 | .8870 | 1.0908 | **30′** |
| 40′ | 829 | 643 | 154 | 243 | 907 | 129 | 857 | 879 | 20′ |
| 50′ | 858 | 669 | 142 | 280 | 894 | 131 | 843 | 850 | 10′ |
| **28° 0′** | .4887 | .4695 | 2.130 | .5317 | 1.881 | 1.133 | .8829 | 1.0821 | **62° 0′** |
| 10′ | 916 | 720 | 118 | 354 | 868 | 134 | 816 | 792 | 50′ |
| 20′ | 945 | 746 | 107 | 392 | 855 | 136 | 802 | 763 | 40′ |
| **30′** | .4974 | .4772 | 2.096 | .5430 | 1.842 | 1.138 | .8788 | 1.0734 | **30′** |
| 40′ | .5003 | 797 | 085 | 467 | 829 | 140 | 774 | 705 | 20′ |
| 50′ | 032 | 823 | 074 | 505 | 816 | 142 | 760 | 676 | 10′ |
| **29° 0′** | .5061 | .4848 | 2.063 | .5543 | 1.804 | 1.143 | .8746 | 1.0647 | **61° 0′** |
| 10′ | 091 | 874 | 052 | 581 | 792 | 145 | 732 | 617 | 50′ |
| 20′ | 120 | 899 | 041 | 619 | 780 | 147 | 718 | 588 | 40′ |
| **30′** | .5149 | .4924 | 2.031 | .5658 | 1.767 | 1.149 | .8704 | 1.0559 | **30′** |
| 40′ | 178 | 950 | 020 | 696 | 756 | 151 | 689 | 530 | 20′ |
| 50′ | 207 | 975 | 010 | 735 | 744 | 153 | 675 | 501 | 10′ |
| **30° 0′** | .5236 | .5000 | 2.000 | .5774 | 1.732 | 1.155 | .8660 | 1.0472 | **60° 0′** |
| 10′ | 265 | 025 | 1.990 | 812 | 720 | 157 | 646 | 443 | 50′ |
| 20′ | 294 | 050 | 980 | 851 | 709 | 159 | 631 | 414 | 40′ |
| **30′** | .5323 | .5075 | 1.970 | .5890 | 1.698 | 1.161 | .8616 | 1.0385 | **30′** |
| 40′ | 352 | 100 | 961 | 930 | 686 | 163 | 601 | 356 | 20′ |
| 50′ | 381 | 125 | 951 | 969 | 675 | 165 | 587 | 327 | 10′ |
| **31° 0′** | .5411 | .5150 | 1.942 | .6009 | 1.664 | 1.167 | .8572 | 1.0297 | **59° 0′** |
| 10′ | 440 | 175 | 932 | 048 | 653 | 169 | 557 | 268 | 50′ |
| 20′ | 469 | 200 | 923 | 088 | 643 | 171 | 542 | 239 | 40′ |
| **30′** | .5498 | .5225 | 1.914 | .6128 | 1.632 | 1.173 | .8526 | 1.0210 | **30′** |
| 40′ | 527 | 250 | 905 | 168 | 621 | 175 | 511 | 181 | 20′ |
| 50′ | 556 | 275 | 896 | 208 | 611 | 177 | 496 | 152 | 10′ |
| **32° 0′** | .5585 | .5299 | 1.887 | .6249 | 1.600 | 1.179 | .8480 | 1.0123 | **58° 0′** |
| 10′ | 614 | 324 | 878 | 289 | 590 | 181 | 465 | 094 | 50′ |
| 20′ | 643 | 348 | 870 | 330 | 580 | 184 | 450 | 065 | 40′ |
| **30′** | .5672 | .5373 | 1.861 | .6371 | 1.570 | 1.186 | .8434 | 1.0036 | **30′** |
| 40′ | 701 | 398 | 853 | 412 | 560 | 188 | 418 | 1.0007 | 20′ |
| 50′ | 730 | 422 | 844 | 453 | 550 | 190 | 403 | 977 | 10′ |
| **33° 0′** | .5760 | .5446 | 1.836 | .6494 | 1.540 | 1.192 | .8387 | .9948 | **57° 0′** |
| 10′ | 789 | 471 | 828 | 536 | 530 | 195 | 371 | 919 | 50′ |
| 20′ | 818 | 495 | 820 | 577 | 520 | 197 | 355 | 890 | 40′ |
| **30′** | .5847 | .5519 | 1.812 | .6619 | 1.511 | 1.199 | .8339 | .9861 | **30′** |
| 40′ | 876 | 544 | 804 | 661 | 501 | 202 | 323 | 832 | 20′ |
| 50′ | 905 | 568 | 796 | 703 | 1.492 | 204 | 307 | 803 | 10′ |
| **34° 0′** | .5934 | .5592 | 1.788 | .6745 | 1.483 | 1.206 | .8290 | .9774 | **56° 0′** |
| 10′ | 963 | 616 | 781 | 787 | 473 | 209 | 274 | 745 | 50′ |
| 20′ | 992 | 640 | 773 | 830 | 464 | 211 | 258 | 716 | 40′ |
| **30′** | .6021 | .5664 | 1.766 | .6873 | 1.455 | 1.213 | .8241 | .9687 | **30′** |
| 40′ | 050 | 688 | 758 | 916 | 446 | 216 | 225 | 657 | 20′ |
| 50′ | 080 | 712 | 751 | 959 | 437 | 218 | 208 | 628 | 10′ |
| **35° 0′** | .6109 | .5736 | 1.743 | .7002 | 1.428 | 1.221 | .8192 | .9599 | **55° 0′** |
| 10′ | 138 | 760 | 736 | 046 | 419 | 223 | 175 | 570 | 50′ |
| 20′ | 167 | 783 | 729 | 089 | 411 | 226 | 158 | 541 | 40′ |
| **30′** | .6196 | 5807 | 1.722 | .7133 | 1.402 | 1.228 | 8141 | .9512 | **30′** |
| 40′ | 225 | 831 | 715 | 177 | 393 | 231 | 124 | 483 | 20′ |
| 50′ | 254 | 854 | 708 | 221 | 385 | 233 | 107 | 454 | 10′ |
| **36° 0′** | .6283 | .5878 | 1.701 | .7265 | 1.376 | 1.236 | .8090 | .9425 | **54° 0′** |
| | | **Cos** | **Sec** | **Cot** | **Tan** | **Csc** | **Sin** | **Radians** | **Degrees** |

## *Table IV. Values of Trigonometric Functions (concluded)*

| Degrees | Radians | Sin | Csc | Tan | Cot | Sec | Cos | | |
|---|---|---|---|---|---|---|---|---|---|
| **36° 0′** | .6283 | .5878 | 1.701 | .7265 | 1.376 | 1.236 | .8090 | .9425 | **54° 0′** |
| 10′ | 312 | 901 | 695 | 310 | 368 | 239 | 073 | 396 | 50′ |
| 20′ | 341 | 925 | 688 | 355 | 360 | 241 | 056 | 367 | 40′ |
| **30′** | .6370 | .5948 | 1.681 | .7400 | 1.351 | 1.244 | .8039 | .9338 | **30′** |
| 40′ | 400 | 972 | 675 | 445 | 343 | 247 | 021 | 308 | 20′ |
| 50′ | 429 | 995 | 668 | 490 | 335 | 249 | 004 | 279 | 10′ |
| **37° 0′** | .6458 | .6018 | 1.662 | .7536 | 1.327 | 1.252 | .7986 | .9250 | **53° 0′** |
| 10′ | 487 | 041 | 655 | 581 | 319 | 255 | 969 | 221 | 50′ |
| 20′ | 516 | 065 | 649 | 627 | 311 | 258 | 951 | 192 | 40′ |
| **30′** | .6545 | .6088 | 1.643 | .7673 | 1.303 | 1.260 | .7934 | .9163 | **30′** |
| 40′ | 574 | 111 | 636 | 720 | 295 | 263 | 916 | 134 | 20′ |
| 50′ | 603 | 134 | 630 | 766 | 288 | 266 | 898 | 105 | 10′ |
| **38° 0′** | .6632 | .6157 | 1.624 | .7813 | 1.280 | 1.269 | .7880 | .9076 | **52° 0′** |
| 10′ | 661 | 180 | 618 | 860 | 272 | 272 | 862 | 047 | 50′ |
| 20′ | 690 | 202 | 612 | 907 | 265 | 275 | 844 | .9018 | 40′ |
| **30′** | .6720 | .6225 | 1.606 | .7954 | 1.257 | 1.278 | .7826 | .8988 | **30′** |
| 40′ | 749 | 248 | 601 | .8002 | 250 | 281 | 808 | 959 | 20′ |
| 50′ | 778 | 271 | 595 | 050 | 242 | 284 | 790 | 930 | 10′ |
| **39° 0′** | .6807 | .6293 | 1.589 | .8098 | 1.235 | 1.287 | .7771 | .8901 | **51° 0′** |
| 10′ | 836 | 316 | 583 | 146 | 228 | 290 | 753 | 872 | 50′ |
| 20′ | 865 | 338 | 578 | 195 | 220 | 293 | 735 | 843 | 40′ |
| **30′** | .6894 | .6361 | 1.572 | .8243 | 1.213 | 1.296 | .7716 | .8814 | **30′** |
| 40′ | 923 | 383 | 567 | 292 | 206 | 299 | 698 | 785 | 20′ |
| 50′ | 952 | 406 | 561 | 342 | 199 | 302 | 679 | 756 | 10′ |
| **40° 0′** | .6981 | .6428 | 1.556 | .8391 | 1.192 | 1.305 | .7660 | .8727 | **50° 0′** |
| 10′ | .7010 | 450 | 550 | 441 | 185 | 309 | 642 | 698 | 50′ |
| 20′ | 039 | 472 | 545 | 491 | 178 | 312 | 623 | 668 | 40′ |
| **30′** | .7069 | .6494 | 1.540 | .8541 | 1.171 | 1.315 | .7604 | .8639 | **30′** |
| 40′ | 098 | 517 | 535 | 591 | 164 | 318 | 585 | 610 | 20′ |
| 50′ | 127 | 539 | 529 | 642 | 157 | 322 | 566 | 581 | 10′ |
| **41° 0′** | .7156 | .6561 | 1.524 | .8693 | 1.150 | 1.325 | .7547 | .8552 | **49° 0′** |
| 10′ | 185 | 583 | 519 | 744 | 144 | 328 | 528 | 523 | 50′ |
| 20′ | 214 | 604 | 514 | 796 | 137 | 332 | 509 | 494 | 40′ |
| **30′** | .7243 | .6626 | 1.509 | .8847 | 1.130 | 1.335 | .7490 | .8465 | **30′** |
| 40′ | 272 | 648 | 504 | 899 | 124 | 339 | 470 | 436 | 20′ |
| 50′ | 301 | 670 | 499 | 952 | 117 | 342 | 451 | 407 | 10′ |
| **42° 0′** | .7330 | .6691 | 1.494 | .9004 | 1.111 | 1.346 | .7431 | .8378 | **48° 0′** |
| 10′ | 359 | 713 | 490 | 057 | 104 | 349 | 412 | 348 | 50′ |
| 20′ | 389 | 734 | 485 | 110 | 098 | 353 | 392 | 319 | 40′ |
| **30′** | .7418 | .6756 | 1.480 | .9163 | 1.091 | 1.356 | .7373 | .8290 | **30′** |
| 40′ | 447 | 777 | 476 | 217 | 085 | 360 | 353 | 261 | 20′ |
| 50′ | 476 | 799 | 471 | 271 | 079 | 364 | 333 | 232 | 10′ |
| **43° 0′** | .7505 | .6820 | 1.466 | .9325 | 1.072 | 1.367 | .7314 | .8203 | **47° 0′** |
| 10′ | 534 | 841 | 462 | 380 | 066 | 371 | 294 | 174 | 50′ |
| 20′ | 563 | 862 | 457 | 435 | 060 | 375 | 274 | 145 | 40′ |
| **30′** | .7592 | .6884 | 1.453 | .9490 | 1.054 | 1.379 | .7254 | .8116 | **30′** |
| 40′ | 621 | 905 | 448 | 545 | 048 | 382 | 234 | 087 | 20′ |
| 50′ | 650 | 926 | 444 | 601 | 042 | 386 | 214 | 058 | 10′ |
| **44° 0′** | .7679 | .6947 | 1.440 | .9657 | 1.036 | 1.390 | .7193 | .8029 | **46° 0′** |
| 10′ | 709 | 967 | 435 | 713 | 030 | 394 | 173 | 999 | 50′ |
| 20′ | 738 | 988 | 431 | 770 | 024 | 398 | 153 | 970 | 40′ |
| **30′** | .7767 | .7009 | 1.427 | .9827 | 1.018 | 1.402 | .7133 | .7941 | **30′** |
| 40′ | 796 | 030 | 423 | 884 | 012 | 406 | 112 | 912 | 20′ |
| 50′ | 825 | 050 | 418 | 942 | 006 | 410 | 092 | 883 | 10′ |
| **45° 0′** | .7854 | .7071 | 1.414 | 1.000 | 1.000 | 1.414 | .7071 | .7854 | **45° 0′** |
| | | **Cos** | **Sec** | **Cot** | **Tan** | **Csc** | **Sin** | **Radians** | **Degrees** |

## *Table V. Degrees and Minutes to Radians*

| Minutes | Radians | Degrees | Radians |
|---|---|---|---|
| 1′ | 0.00029 | 1° | 0.01745 |
| 2′ | 0.00058 | 2° | 0.03491 |
| 3′ | 0.00087 | 3° | 0.05236 |
| 4′ | 0.00116 | 4° | 0.06981 |
| 5′ | 0.00145 | 5° | 0.08727 |
| 6′ | 0.00175 | 6° | 0.10472 |
| 7′ | 0.00204 | 7° | 0.12217 |
| 8′ | 0.00233 | 8° | 0.13963 |
| 9′ | 0.00262 | 9° | 0.15708 |
| 10′ | 0.00291 | 10° | 0.17453 |
| 15′ | 0.00436 | 15° | 0.26180 |
| 30′ | 0.00873 | 30° | 0.52360 |
| 45′ | 0.01309 | 60° | 1.04720 |
| 60′ | 0.01745 | 90° | 1.57080 |

1″ = 0.0000048 radian
60″ = 1′ = 0.0002909 radian
3600″ = 60′ = 1° = 0.01745329 radian
180° = $\pi$ radians = 3.14159265 radians

## *Table VI. Radians to Degrees and Minutes*

| Radians | Degrees and Minutes | Radians | Degrees and Minutes |
|---|---|---|---|
| 0.001 | 0° 3.44′ | 0.1 | 5° 43.77′ |
| 0.002 | 0° 6.88′ | 0.2 | 11° 27.55′ |
| 0.003 | 0° 10.31′ | 0.3 | 17° 11.32′ |
| 0.004 | 0° 13.75′ | 0.4 | 22° 55.10′ |
| 0.005 | 0° 17.19′ | 0.5 | 28° 38.87′ |
| 0.006 | 0° 20.63′ | 0.6 | 34° 22.65′ |
| 0.007 | 0° 24.06′ | 0.7 | 40° 6.42′ |
| 0.008 | 0° 27.50′ | 0.8 | 45° 50.20′ |
| 0.009 | 0° 30.96′ | 0.9 | 51° 33.97′ |
| 0.01 | 0° 34.38′ | 1.0 | 57° 17.75′ |
| 0.02 | 1° 8.75′ | 2.0 | 114° 35.49′ |
| 0.03 | 1° 43.13′ | 3.0 | 171° 53.24′ |
| 0.04 | 2° 17.51′ | 4.0 | 229° 10.99′ |
| 0.05 | 2° 51.89′ | 5.0 | 286° 28.73′ |
| 0.06 | 3° 26.26′ | 6.0 | 343° 46.48′ |
| 0.07 | 4° 0.64′ | 7.0 | 401° 4.25′ |
| 0.08 | 4° 35.02′ | 8.0 | 458° 21.97′ |
| 0.09 | 5° 9.40′ | 9.0 | 515° 30.72′ |

# ANSWERS TO ODD-NUMBERED PROBLEMS

## CHAPTER I

### § 1-2

**1.** (a) $13\sqrt{5}/2$. (b) 5. (c) $2\sqrt{17}$. (d) $4\sqrt{2}$. **7.** Yes. **9.** Yes. **11.** (5, 3), (8, 5). **17.** $|a + b| \leq |a| + |b|$ always; $|a + b| < |a| + |b|$ if and only if $a$ and $b$ are of different sign.

### § 1-3

**1.** On line in (a), (c), (d). **3.** Rectangle in (d), (f); parallelogram but not rectangle in (a), (c). **5.** 81°12′, 42°19′, 56°29′. **7.** $m_1 = -m_2$. **9.** (−1, 1), (−5, −13). **11.** (a) (6, 4), (−2, −2). (b) $(\frac{42}{5}, \frac{29}{5})$, $(-\frac{22}{5}, -\frac{19}{5})$. **13.** (3, −1) or (1, 3). **15.** $-3, \frac{1}{3}$. **17.** $a^2 = b^2 + c^2$. All sides equal.

### § 1-4

**1.** (a) $2x + y = 5$; (b) $4x - y = -4$. (c) $y = 3x + 6$; (d) $2x - 5y + 24 = 0$; (e) $5x - 2y = 34$; (f) $3x + y = 13$; (g) $x + 2y = 4$; (h) $x + y + 1 = 0$; (i) $5x - 8y + 40 = 0$. **5.** (a) $x - \sqrt{3}y + 2 + \sqrt{3} = 0$; (b) $3x + 4y + 25 = 0$; (c) $6x - 4y = 19$; (d) $3x - y = 2$; (e) $4x - 3y + 18 = 0$; (f) $2x - 3y = 6$; (g) $3x + y = 8$. **7.** (a) $5x - 9y = 160$. (b) Slope $\frac{5}{9}$. (c) 0°F = −17.77 ··· °C. (d) −40. (e) 37°C. (f) 64.4°F. **9.** $(\frac{53}{14}, -\frac{31}{14})$.

### § 1-5

**1.** (a) $x^2 + y^2 = 2x$; (b) $x^2 + y^2 = 2(x + y)$; (c) $x^2 + y^2 + 4x - 6y + 9 = 0$; (d) $x^2 + y^2 - 6x - 4y + 4 = 0$; (e) $(x - 6)^2 + (y \pm 4)^2 = 16$. **3.** (a) $x^2 = -2y$.

§ 1-6

**1.** $P = 2x + \dfrac{8}{x}\cdot$ **3.** $A = 100x - \dfrac{x^2}{2}\cdot$ **5.** $y = \sqrt{100 - x^2}$, $A = 4x\sqrt{100 - x^2}$. **7.** $V = \dfrac{\pi y}{4}(64 - y^2)$, $S = \pi y\sqrt{64 - y^2} + \dfrac{\pi}{2}(64 - y^2)$. **11.** No. **13.** $D = \sqrt{1 + x^2}$ if $x < 0$; $D = 1$ if $0 \leq x \leq 1$; $D = \sqrt{x^2 - 2x + 2}$ if $1 < x$. **15.** (a) $x = a - \dfrac{12a}{\sqrt{36 + a^2}}\cdot$

§ 1-7

**1.** (a) $v = 96 - \frac{3}{2}t^2$, $a = -3t$, $v > 0$ if $|t| < 8$, $v < 0$ if $|t| > 8$. $v$ decreasing if $t > 0$. (b) 400. No. (c) $s = -16$, $v = -12$. $1 < t < 5$. (d) $s$ increases if $0 < t < 2$, decreases if $2 < t$. $v$ increases if $0 < t < 1$. (e) Max. $s$ is 64. $a = 128$ at $t = 0$, $a = -256$ at $t = \sqrt{2}$. (f) $v = 0$: $t = 1$, $s = 32$ and $t = 3$, $s = 0$. Before $t = 1$, $s$ increasing; $s$ decreasing if $1 < t < 3$. $v$ decreasing if $t < 2$, increasing if $t > 2$. **3.** 16. **5.** $\frac{5}{2}$. **7.** $18\pi t$. **9.** (a) $2y/\sqrt{3}$; (b) $\sqrt{3}x/2$. **11.** (a) $x = 1, 5$; (b) $x = -4, 2$; (c) $x = -1, 2$; (d) $x = 1, -2$; (e) $x = -1$; (f) $x = 2, \dfrac{-1 \pm \sqrt{17}}{4}\cdot$ **13.** $T$ decreases $(29/14{,}500)^\circ$ F for each additional foot above sea level. **15.** $\dfrac{dF}{dC} = \dfrac{9}{5}, \dfrac{dC}{dF} = \dfrac{5}{9}\cdot$ **17.** (a) 9000; (b) 600. **19.** $-\frac{1}{4}$.

§ 1-8

**1.** $-26, -2, \dfrac{3q}{p} - \dfrac{p^3}{q^3}, 3x - x^{-3}$. **3.** (1) (a) None. (b) 0. (c) 0, −8. (d) 4, 5. (e) $-1 \leq x \leq 1$. (f) $-2 < x \leq 0$. (g) $\pm 3$. (h) −3, 0, 2. **5.** (a) $(2 - x)^{-2}$. (b) $-12x^{-5}$. (c) $\dfrac{-2}{(1 + x)^2}\cdot$ (d) $\dfrac{1 - x^2}{(x^2 + 1)^2}\cdot$ (e) $\dfrac{1 - 2x}{x^2(x - 1)^2}\cdot$ (f) $\dfrac{x(x + 4)}{(x + 2)^2}\cdot$ **7.** (a) $[f(x)]^n = f(nx)$. (b) $f(x + y) = f(x) + f(y)$. **9.** $f[g(x)] = x$ for all $x$. $g[f(x)] = x$ if $x > 0$. **11.** (a) $\frac{1}{2}$, (b) $-\frac{3}{2}$. (c) $2/3a$. (d) $-3/4a$. (e) 0. (f) $1/2a$. (g) $1/2\sqrt{a}$. (h) $1/2|a|$. **13.** (a) $\frac{1}{5}$. **17.** (a) $x$ less than $10^{-4}$, $10^{-4}/81$, $k^2$, respectively.

§ 1-9

**1.** (a) $3x - y = 4$. (b) $72x - 16y = 81$. (c) $15x + 2y + 4 = 0$ at $x = -\frac{1}{2}$; $y = 0$ at $x = 0$; $6x - y = 2$ at $x = 1$. (d) $y = 96x + 256$ at $x = 0$; $y = 400$ at $x = 3$; $y = 1280 - 160x$ at $x = 8$. (e) $y = 96x$ at $x = 0$; $y = 512$ at $x = 8$; $y = -198x + 2744$ at $x = 14$. (a) $32x - 16y = 5$. (b) $72x - 16y + 85 = 0$ at $x = -\frac{9}{4}$; $8x + 16y = 5$ at $x = \frac{1}{4}$; $20x + 40y = 13$ at $(\frac{1}{4}, \frac{1}{5})$. (c) $\tan^{-1}\frac{1}{4}$ and $\pi/2$. **5.** $-4 \pm \sqrt{7}$. **11.** (a) Orth. at $x = \frac{1}{2}$. (b) Not orth. at $x = \frac{1}{2}$. (c) Not orth. at $x = \frac{1}{2}$. (d) Orth. at $x = \pm 1$. (e) Not orth. at

$x = \pm 1$. (f) Orth. at $x = \pm 1\sqrt{2}$. (g) Orth. at $x = \pm 1$. (h) Orth. at $x = \pm 2\sqrt{2}$. (i) Not orth. at $x = \pm 1$.

### § 1-10

**1.** Critical values of $x$ are listed. (a) 0, 6. (b) $\pm 1$. (c) $\pm 3$. (d) 1, 2. (e) 2. (f) 1, 3. (g) 0, $\pm 2$. (h) 0, $-3$. (i) 1, $-2$. (j) 0, 3. (k) 0, 2, $-\frac{1}{2}$. (l) 1, 3. (m) 0, 1. (n) 0, 1, 2, $-2$. **3.** $V = \frac{5\pi}{2} x^2(6 - x)$. **5.** $V = 4x(12 - x)(8 - x)$. **7.** $x = \frac{47}{12}$. **9.** $x = \frac{4}{3}$ (approximate).

## CHAPTER II

### § 2-1

**1.** (a) $x = 1$ not included. (b) $f$ not continuous at $x = 1$. (c) $f$ not continuous at $x = 0$. (d) $x$ not confined to a finite interval.

### § 2-2

**3.** (a) $y = \frac{1}{3}x^3 - \frac{1}{2}x^2 + 1$. (b) $y = 2 - x^4$. (c) $y = \frac{1}{4}x^4 - x^3 + x^2 + 4$. **5.** (a) $y = -\frac{3}{2}x^2 + x + \frac{3}{2}$. (b) $y = \frac{1}{3}x^3 - 6x + 9$. (c) $y = -\frac{1}{12}x^4 + \frac{1}{6}x^3 + \frac{8}{3}x + \frac{2}{3}$. (d) $y = -x^5 + 18x + 2$.

### § 2-3

**1.** (a) $s = -4.9t^2 + 196t$. (b) 1960 m in 20 sec. **3.** (a) $s = -4.9t^2 + 30t + 50$. (b) About $7\frac{1}{2}$ sec. **5.** (a) 99 ft. (b) 143 ft. **7.** 49 m/sec. 122.5 m. **9.** $s = 12t^2 + 49t - 135$. **11.** 40 ft/sec. **13.** (a) 11 sec. (b) 484 ft. **15.** (a) 42 ft. (b) 3 in. **17.** (a) 1, 8. (b) 112 ft/sec. (c) 196 ft in $3\frac{1}{2}$ sec. **19.** 3 yd/sec² (approximate).

### § 2-4

**1.** (a) $x^2 + 16 = 8y$; (b) $x^2 + 6y = 9$; (c) $y^2 + 4x + 4 = 0$; (d) $y^2 = 12(x - 1)$; (e) $y^2 = 8x$; (f) $x^2 = 12(y + 3)$; (g) $x^2 = 8y$; (h) $y^2 + 16x = 64$. **3.** (a) $(-2, 0)$, $(-\frac{3}{2}, 0)$, $x = -\frac{5}{2}$. (b) $(1, -\frac{1}{4})$, $(1, \frac{3}{4})$, $y = -\frac{5}{4}$. (c) $(-1, 2)$, $(-1, 3)$, $y = 1$. (d) $(4, 3)$, $(-2, 3)$, $x = 10$. (e) $(-4, -2)$, $(-4, 2)$, $y = -6$. (f) $(\frac{9}{2}, 1)$, $(2, 1)$, $x = 7$. (g) $(-2, 3)$, $(-2, 1)$, $y = 5$. (h) $(6, -1)$, $(6, -4)$, $y = 2$. (i) $(-4, 5)$, $(-4, \frac{5}{2})$, $y = \frac{15}{2}$. (j) $(4, -1)$, $(\frac{23}{4}, -1)$, $x = \frac{9}{4}$. (k) $(\frac{7}{6}, -4)$ $(\frac{25}{6}, -4)$, $x = -\frac{11}{6}$. (l) $(1, 2)$, $(1, \frac{7}{3})$, $y = \frac{5}{3}$. (m) $(-2, 3)$, $(-\frac{35}{16}, 3)$, $x = -\frac{29}{16}$. (n) $(-2, 1)$, $(-\frac{29}{12}, 1)$, $x = -\frac{19}{12}$. **5.** (a) $8x^2 + 9y = 72$. (b) $16y = 5x^2$. (c) $5y^2 = 9(x + 1)$. (d) $3(y - 4)^2 = -8(x - 6)$. **7.** $y = x^2 - x$. **9.** (a) $15y = -4x^2 + 31x + 3$. (b) $126y = 11x^2 - 63x + 52$. (c) $3y = -4x^2 + 26x - 33$. (d) $12y = 7x^2 - 9x - 34$. **11.** $\frac{63}{4}$, 27, $\frac{135}{4}$. **13.** 9 in. **15.** (a) 4 ft. (b) 30 ft. **17.** Greatest height is 8 in., at the 20 ft station. Heights at stations, in order: 0, $3\frac{1}{2}$, 6, $7\frac{1}{2}$, 8, $7\frac{1}{2}$, 6.

§ 2-5

**1.** (a) $y = 2x - 4$; (b) $y = 4(x - 1)$; (c) $2x - 3y = 3$; (d) $y = 1 - x$; (e) $y = 3 - 2x$; (f) $4x + 3y = 12$. **3.** (a) $2x - 3y = 3$; (b) $10x - y = 100$; (c) $2x - y = 8$; (d) $y = \pm 2x - 6$.

§ 2-6

**1.** $\frac{9}{2}$.

§ 2-7

**3.** (a) 18. (b) 16. (c) $\frac{7}{3}$. (d) $\frac{8}{3}$. (e) $\frac{27}{4}$. (f) $\frac{65}{3}$. **7.** (a) 54. (b) 32. (c) 10. **9.** (a) 0, 32, 16, 0. (b) $x = 1$. (c) $x = 2$.

## *Review Problems, End of Chapter II*

**1.** $6x - 4y = 9$, $6x + 9y = -4$. $(\frac{5}{6}, -1)$. **3.** Focus (0, 0). **5.** $30\pi$. **7.** 144. **9.** No. For $A > 24$. One real root if $A = 24$. Three real roots, one of them double, if $A = -3$. **13.** (a) $y = 2x^4 - x^2 - 20$. (b) $y = \frac{1}{3}x^3 - \frac{1}{2}x^2 - \frac{7}{2}$. **17.** $s = -9t^2 + 72t - 64$. Max. $s = 80$.

## CHAPTER III

§ 3-2

**1.** (a) $3(5x^4 + 19x^2 + 2)$.
(b) $-5x^4 + 12x^3 + 16x - 24$.
(c) $x(x - 1)(5x^2 + 5x + 2)$.
(d) $3(x + 1)(x - 3)$.
(e) $2(x - 3)(x^2 - 3x - 1)$.
(f) $(x + 1)(5x^3 - 5x^2 - 4x + 2)$.
(g) $x(x - 1)(7x^4 + 7x^3 + 2x^2 - 2x - 2)$.
(h) $3x^2 + 6x + 2$.
(i) $x(x + 2)(7x^4 - 14x^3 + 8x^2 + 16x - 32)$.
(j) $9x(x - 3)^2(x + 3)(x^4 + 3x^3 + 11x^2 + 15x + 18)$.
(k) $2x^3(x^2 - 4)(5x^4 + 4x^2 - 32)$.
(l) $2(3x + 2)^2(243x^3 - 189x^2 + 36x - 12)$.

**3.** (a) $-\dfrac{36x}{(x^2 - 9)^2}$. (b) $-\dfrac{6}{(2x - 5)^2}$. (c) $\dfrac{3x^2}{(8 - x^3)^2}$.

(d) $-\dfrac{100x}{(x^2 + 25)^2}$. (e) $\dfrac{5 - 4x^2}{(5 + 4x^2)^2}$. (f) $\dfrac{8 - 16x^3 - x^6}{(x^3 + 2)^2}$.

(g) $-\dfrac{x^4 - 44x^2 + 64}{(16 - x^2)^2}$. (h) $\dfrac{x^2 - 8x + 14}{(x - 4)^2}$.

(i) $\dfrac{2x^2 - 6x - 1}{(2x - 3)^2}$. (j) $\dfrac{x^4(3x^2 + 5)}{(x^2 + 1)^2}$.

(k) $\dfrac{16x}{(x^2+4)^2}$. (l) $\dfrac{2x^2(3a-x)}{(2a-x)^2}$.

§ 3-3

**1.** (a) $12x^2(2x^3-3)$. (b) $42x(7x^2-5)^2$.
(c) $10x^4(2x^3-1)(x^3-2)^4$. (d) $4(3x-2)(3x-1)(x-1)$.
(e) $\dfrac{4x}{(36-x^2)^3}$. (f) $-\dfrac{8x}{(4x^2+9)^2}$.
(g) $-3x^6(4x-9)(3-2x)^2(32x^2-96x+63)$.
(h) $12x(x^3+8)(x^2-4)^2(x^3-2x+4)$.
(i) $\dfrac{2x(16+x^2)(48-x^2)}{(16-x^2)^2}$.
(j) $\dfrac{18x^3-33x^2+12x+2}{(1-3x)^4}$.

**3.** (a) $2v(v^2-25)^2(4v^2-25)$.
(b) $-2(2v+3)^2(5v^2+3v+3)$.
(c) $\dfrac{v(1+3v)(9v^2+13v+2)}{(v+1)^2}$.
(d) $\dfrac{8v(v^2+12)}{(v^2-4)^3}$.
(e) $\dfrac{12v(v^2-4)^2(v^2+4)}{(3v^2+4)^2}$.
(f) $\dfrac{4v(14-v^2)}{(v^2+25)^4}$.
(g) $\dfrac{2(4v^4-8v^3+4v-1)}{(1-2v)^5}$.
(h) $\dfrac{v(v-2a)(3v^2-6av+4a^2)}{(v-a)^2}$.

**5.** $-\frac{1}{2}$ unit/min. **7.** $\frac{5}{2}$ ft/sec.

§ 3-4

**1.** (a) Concave upward if $x < 0$ or $3 < x$, downward if $0 < x < 3$. (b) Concave upward if $x > 0$, downward if $x < 0$. (c) Concave upward if $x < -\frac{1}{2}$ or $\frac{1}{2} < x$, downward if $-\frac{1}{2} < x < \frac{1}{2}$. (d) Concave upward if $x > 1$, downward if $x < 1$. (e) Concave upward if $x > 2$ or $x < 0$, downward if $0 < x < 2$. (f) Concave upward if $x > -2$ or $x < -3$, downward if $-3 < x < -2$.

**3.** (a) Concave upward if $|x| > \dfrac{a}{\sqrt{3}}$, downward if $|x| < \dfrac{a}{\sqrt{3}}$. (b) Concave upward if $x > a\sqrt{3}$ or $-a\sqrt{3} < x < 0$, downward if $0 < x < a\sqrt{3}$ or $x < -a\sqrt{3}$. (c) Only point of inflection at $x = 0$. Concave upward if $0 < x < 1$ or $x < -1$, downward if $-1 < x < 0$ or $1 < x$. (d) Concave upward if $x^2 > 4$, downward if $x^2 < 4$. (e) Only point of inflection at $x = 0$. Concave upward if $x > 0$, downward if $x < 0$.

§ 3-5

**1.** Asymptotes are listed. (a) $x = 2$, $y = 0$. (b) $x = -1$, $y = 0$. (c) $x = \pm 2$, $y = 1$. (d) $x = 2$, $y = 1$. (e) $x = 1$, $y = 1$. (f) $x = \pm 3$, $y = 0$. (g) $x = 0$, $x = 3$, $y = 0$. (h) $x = -1$, $x = 2$, $y = -1$. (i) $x = 0$, $y = 0$. (j) $x = 2$, $y = 0$.

§ 3-6

**1.** (a) $-\dfrac{16}{3}(1-x)^{1/3} + \dfrac{4(3x-1)}{3(3x^2-2x+1)^{2/3}} + \dfrac{2(1-x)}{3(2x-x^2)^{4/3}}$.

(b) $5(2x-1)^{3/2} + \dfrac{3x^2 - x + 1}{(2x^3 - x^2 + 2x - 1)^{3/2}}$.

(c) $-\dfrac{3}{(1+x)^2}\left(\dfrac{1-x}{1+x}\right)^{1/2}$. (d) $\dfrac{2(x^2-16)}{3x^{5/3}(x^2+16)^{1/3}}$.

(e) $-(4 - x^{2/3})^{1/2}x^{-1/3}$. (f) $\dfrac{2x^2 + x - 4}{(x+2)^{3/4}(x-2)^{1/4}}$.

**3.** (a) $\dfrac{13\pi}{96\sqrt{6}}$. (b) Concave downward.

§ 3-7

**1.** (a) $y^2(y + 3xy')$. (b) $\dfrac{xy' + y}{2\sqrt{xy}}$. (c) $\dfrac{2y(xy' - y)}{x^3}$.

(d) $\dfrac{y(1+3xy) - x(xy+3)y'}{2x^{1/2}y^{5/2}}$.

**3.** (a) $\dfrac{8x}{y(3y-8)}$. (b) $\dfrac{-xy}{x^2 + 2y^2}$.

(c) $-y^{1/3}x^{-1/3}$. (d) $-\sqrt{y/x}$.

§ 3-8

**1.** (a) Circle, center $(-1, 3)$, $r = 2$. (b) Ellipse, center $(2, -2)$, foci $(2, -2 \pm \sqrt{5})$, major axis 6, minor axis 4. (c) Point $(2, -1)$. (d) Circle, center $(-2, -1)$, $r = 1$. (e) Ellipse, center $(0, 1)$, foci $(\pm 4, 1)$, major axis 10, minor axis 6. (f) Point $(5, -3)$. (g) Ellipse, center $(\frac{25}{4}, 0)$, foci $\left(\dfrac{25 \pm 9}{4}, 0\right)$, major axis $\frac{15}{2}$, minor axis 6. (h) Circle, center $(\frac{3}{4}, -\frac{2}{3})$, $r = \frac{5}{4}$. (i) Circle, center $(-7, 5)$,

$r = 8$. (j) Ellipse, center $(-1, 2)$, foci $\left(-1, 2 \pm \frac{\sqrt{65}}{6}\right)$, major axis $\sqrt{13}$, minor axis $\frac{2}{3}\sqrt{13}$.

**3.** (a) $x^2 + y^2 + 4x - 6y = 21$.
(b) $x^2 + y^2 - 8(x + y) + 16 = 0$.
(c) $x^2 + y^2 - 12x + 11 = 0$.
(d) $x^2 + y^2 - 12x - 2y = 132$.
(e) $x^2 + y^2 - 4x - 6y + 9 = 0$.
(f) Center at $(1 + \frac{1}{5}\sqrt{155}, 1 - \frac{2}{5}\sqrt{155})$.
(g) $x^2 + y^2 - 8x - 4y = 25$.
(h) $3x^2 + 3y^2 + 3x - 11y = 0$.
(i) $x^2 + y^2 - 12x - 18y + 92 = 0$.

**5.** (a) $5x^2 + 9y^2 = 180$. (b) $89x^2 + 64y^2 = 5696$.
(c) $25x^2 + 16y^2 = 100$. (d) $21x^2 + 25y^2 = 525$.
(e) $4x^2 + 3y^2 = 192$. (f) $99x^2 + 100y^2 = 40{,}000$.

**7.** $(-4, 3), (3, 4)$. **9.** Circle, center $(4, \frac{5}{3})$, $r = \frac{2}{3}\sqrt{37}$. **11.** Circle, center $(-3, 0)$, $r = 8$. **13.** (a) $x^2 + 4y^2 = 20$. (b) $3x^2 + y^2 = 28$. **15.** $25(x^2 + y^2) - 54x = 319$. **19.** Inside the circle $(x - 54)^2 + (y + 32)^2 = 3600$.

§ 3-9

**3.** (a) $16x^2 - 9y^2 = 576$. (b) $16x^2 - 25y^2 = 400$.
(c) $3y^2 - x^2 = 12$. (d) $8x^2 - y^2 = 32$.
(e) $x^2 - 4y^2 = 9$. (f) $12y^2 - 4x^2 = 27$.

**5.** (a) $8x + y = 15$. (b) $13x - 5y = 25$.
(c) $16x - 5y = 54$. (d) $5x - 2y = 16$.
(e) $8x - 25y + 58 = 0$. (f) $6x - 5y = 17$.
(g) $2x - y + 7 = 0$. (h) $4x - y = 32$.

**7.** Right-hand branch of $\frac{x^2}{a^2} - \frac{y^2}{c^2 - a^2} = 1$ (assuming $c > a > 0$).

**9.** (a) $x = \pm\frac{4}{13}\sqrt{13}$, $y = \pm\frac{4}{13}\sqrt{39}$.
(b) $x = \pm\sqrt{6}$, $y = \pm\sqrt{6}$.

**13.** (b) $x^2 + \left(y - \frac{c^2}{a}\right)^2 = a^2 + c^2$.

§ 3-10

**1.** $4 \times 4 \times 2$. **3.** $\frac{1}{8}$ sq mile. **5.** $4 = \frac{8}{3} + \frac{4}{3}$. **7.** $12 = 8 + 4$. **9.** $x = 5$. **11.** (a) $BP = 3\sqrt{2}$. (b) Direct, $C$ to $A$. (c) $r = \frac{13}{5}$. **13.** $8\sqrt{2} \times 4\sqrt{2}$. **15.** 15. **17.** $D = 2H/3$.

§ 3-11

**1.** 2. **5.** $(3 - \sqrt{3})/2$. **7.** 8. **9.** $3a/2$. **13.** (a) $(b/a)^{3/2}$. (b) $a + b$.

§ 3-12

**1.** 1.2 ft/sec. **3.** $\frac{20}{3}$ miles/min. **5.** $-\frac{44}{9}$ cm/sec. **7.** $-\frac{32}{75}$ ft/min. **9.** $\frac{50}{3}$ ft/sec.

## CHAPTER IV

§ 4-2

**1.** (a) $10 \cos (5x - 7)$. (b) $-10 \sin (2x - 3)$.

(c) $\dfrac{\cos \sqrt{x}}{2\sqrt{x}}$. (d) $-12(3x - 4) \sin 2(3x - 4)^2$.

(e) $-3 \sin x \cos x(\cos x + \sin x)$.

(f) $x(3x \cos 2x - 2x^2 \sin 2x - 2 \sin 3x - 3x \cos 3x)$.

(g) $\cos x\,(\cos^2 x - 2 \sin^2 x)$.

(h) $2 \sin^2 2x \cos 2x\,(3 \cos^2 2x - 2 \sin^2 2x)$.

(i) $2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x}$.

(j) $2x \left(2x^2 \sin \dfrac{1}{x^2} - \cos \dfrac{1}{x^2}\right)$.

(k) $-\dfrac{3 \sin 6x}{\sqrt{\cos 6x}}$. (l) $-18 \sin 8x\sqrt{3 \cos^2 4x + 1}$.

§ 4-3

**1.** (a) $6 \tan^2 2x \sec^2 2x$; $24 \tan 2x \sec^2 2x[\tan^2 2x + \sec^2 2x]$.

(b) $3x^2 \sec^2 x^3$; $6x \sec^2 x^3\,(1 + 3x^2 \tan x^3)$.

(c) $10 \sec^2 5x \tan 5x$; $50 \sec^2 5x(2 \tan^2 5x + \sec^2 5x)$.

(d) $\dfrac{8}{x^2} \csc^2 \dfrac{8}{x}$; $\dfrac{16}{x^3} \csc^2 \dfrac{8}{x}\left(\dfrac{8}{x} \operatorname{ctn} \dfrac{8}{x} - 1\right)$.

(e) $-\dfrac{3 \csc 3x \operatorname{ctn} 3x}{x} - \dfrac{\csc 3x}{x^2}$; $\dfrac{9 \csc 3x}{x}(\csc^2 3x + \operatorname{ctn}^2 3x)$

$$+ \frac{6}{x^2} \csc 3x \operatorname{ctn} 3x + \frac{2}{x^3} \csc 3x.$$

(f) $\operatorname{ctn}^2 2x - 4x \operatorname{ctn} 2x \csc^2 2x$;

$$-8 \csc^2 2x(\operatorname{ctn} 2x - 2x \operatorname{ctn}^2 2x - x \csc^2 2x).$$

(g) $\dfrac{3 \sec^2 3x}{2\sqrt{1 + \tan 3x}}$; $\dfrac{9 \sec^2 3x[3 \tan^2 3x + 4 \tan 3x - 1]}{4(1 + \tan 3x)^{3/2}}$.

(h) $\dfrac{1}{(x - 1)^2} \csc^2 \dfrac{x}{x - 1}$; $\dfrac{2}{(x - 1)^3} \csc^2 \dfrac{x}{x - 1}\left(\dfrac{1}{x - 1} \operatorname{ctn} \dfrac{x}{x - 1} - 1\right)$.

**3.** $40\pi\sqrt{3}$ ft/min. **5.** $10\pi$ miles/min. **7.** $25\pi\sqrt{3}/2$ ft/min. $25\pi/2$ ft/min. **9.** $\pi/3$. No.

§ 4-4

**1.** (a) $\pi/4$; (b) $2\pi/3$; (c) $\pi/3$; (d) $5\pi/12$; (e) $\pi/12$. **3.** $\cos^{-1} x + \cos^{-1}(-x) = \pi$. 2.1991.

**5.** (a) $\frac{1}{2\sqrt{x - x^2}}$; (b) $\frac{-1}{\sqrt{4x - x^2}}$; (c) $\frac{-2}{\sqrt{9x - 4x^2}}$;

(d) $\frac{-2}{1 + x^2}$; (e) $\frac{1}{1 + x^2}$; (f) $\frac{2}{\sqrt{2 - 4x - 4x^2}}$;

(g) $\frac{1}{1 + x^2}$; (h) $\frac{6}{9 + x^2}$.

**7.** (a) $-\frac{1}{2}$; (b) $\pi/2$; (c) $\pi/2$, $-\pi/2$. **9.** $\theta = \operatorname{ctn}^{-1}\frac{x}{b} - \operatorname{ctn}^{-1}\frac{x}{a}$. $x = \sqrt{ab}$. **11.** (a) $\frac{2x}{|x|(1 + x^2)}$; (b) $\frac{-x}{|x|\sqrt{1 - x^2}}$. **13.** $\left(1, \frac{\pi}{2}\right)$. **15.** (a) $y = \sqrt{1 - x^2}$; (b) $y = \frac{\sqrt{1 - x^2}}{x}$, (c) $y = \frac{x}{\sqrt{1 + x^2}}$; (d) $y = \frac{1}{x}$.

§ 4-5

**1.** (a) $2\pi b^2$. (b) $4\pi b^3/3\sqrt{3}$. **3.** $\sqrt{3}$. **5.** (a) $\cos^{-1}\frac{1}{3}$. (b) $288\pi$. **7.** $1\frac{5}{18}$ hr. **9.** $(a^{2/3} + b^{2/3})^{3/2}$. **11.** (a) Max. at $\pi/3$; min. at $5\pi/3$. (b) Max. at $\pi/3$; min. at $2\pi/3$. (c) Max. at 0, $2\pi$; min. at $\pi$. (d) Max. at $\pi/3$, $5\pi/3$; min. at 0, $\pi$, $2\pi$. (e) Max. at 0, $2\pi$; min. at $\pi$. (f) Max. at 0, $\pi \pm \cos^{-1}\frac{1}{\sqrt{5}}$, $2\pi$; min. at $\cos^{-1}\frac{1}{\sqrt{5}}$, $\pi$, $2\pi - \cos^{-1}\frac{1}{\sqrt{5}}$. (g) Max. at $\pi/18$, $5\pi/18$; min. at $\pi/6$, $\pi/2$. **13.** 100 ft. **15.** (a) $2\sin^{-1}\frac{b}{2c}$. (b) $-1.2$ radians/min. (c) 20 ft/min. **17.** $\frac{126}{169}$ radian/sec.

§ 4-6

**1.** (a) $5\sqrt{17}$ ft., $2\pi$sec. (b) $\sin^{-1}\frac{3}{\sqrt{17}}$ sec. (c) $\cos^{-1}\frac{1}{\sqrt{17}}$ sec. **3.** (a) $60/\pi$ miles. (b) 60 miles/hr. (c) 30 miles/hr. $30\sqrt{3}/\pi$ miles. (d) $x = \frac{30}{\pi}\cos \pi t + \frac{30\sqrt{3}}{\pi}\sin \pi t$. (e) 0, $-60$. **5.** (a) $\frac{1}{12}$ the period. (b) $\frac{1}{6}$ the period. **7.** (a) 13 ft. (b) 5 ft. (c) $x = 5\cos 6\pi t - 12\sin 6\pi t$. **9.** (a) $x = 2\cos\frac{t}{2} - 6\sin\frac{t}{2}$. (b) $2\sqrt{10}$. (c) 0.6434.

## *Review Problems, End of Chapter IV*

**3.** (a) $S = \pi(4a^2h^2 - 2ah^3)^{1/2}$. (b) $8\pi a^2/3\sqrt{3}$. (c) No. **7.** Min. $\pi/4$ hr.; max. $\dfrac{\pi + 6\sqrt{3}}{12}$ hr. **9.** 96 ft/sec. **11.** Slope is $K$. **13.** (a) $\frac{4}{3}$ ft/sec. (b) $\frac{16}{3}$ ft/sec. (c) $\frac{20}{3}$ ft/sec. **17.** $\sin^{-1}\frac{1}{3}$. **19.** (a) $\dfrac{72}{\sqrt{9 - 4t^2}}$.

# CHAPTER V

### § 5-2

**1.** (a) $a^2(a^2 + x^2)^{-3/2}\,dx$. (b) $-2\csc^2 2x\,dx$. (c) $\dfrac{2x\,dx}{\sqrt{1 - x^4}}$. (d) $\dfrac{dx}{2x\sqrt{2x - 1}}$. (e) $\cos^{-1} x\,dx$. (f) $4\sin^2 2x\,dx$. (g) $12\tan 3x\sec^2 3x\,dx$.

**3.** (a) $\dfrac{8x^2}{3y}$. (b) $\dfrac{2x - y}{x - 2y}$. (c) $-\dfrac{x^2 + y^2}{2xy + y^2}$.

(d) $\dfrac{8y(y - x)}{(2x - 3y)(2x - 5y)}$. (e) $\dfrac{\sin x}{\cos y}$.

(f) $\dfrac{y - 2x\tan^{-1}\dfrac{y}{x}}{x + 2y\tan^{-1}\dfrac{y}{x}}$. (g) $\dfrac{-\sin^2 y}{27\cos y + 64}$.

(h) $\dfrac{x^2\sin x\cos x + x\sin^2 x - xy^4}{y + 2x^2y^3}$.

**5.** $\dfrac{dx}{5 + 3\sin x}$.

### § 5-4 (A Constant $C$ should be added in each case.)

**1.** (a) $-\frac{1}{3}(1 - 2x)^{3/2}$. (b) $-\frac{1}{5}\cos 5x$.
(c) $-\frac{2}{3}(2 - 3x)^{1/2}$. (d) $(9 - x^2)^{-1/2}$. (e) $\frac{1}{6}\sin^3 2x$.

(f) $-2\sin\left(\dfrac{1 - x}{2}\right)$. (g) $-\frac{4}{3}\left(\cos\dfrac{x}{2}\right)^{3/2}$.

(h) $\frac{1}{4}\tan 4x$. (i) $-\frac{1}{3}\csc 3x$. (j) $\operatorname{ctn}(2 - u)$.
(k) $-\frac{1}{3}\sin^{-1}(1 - 3y)$. (l) $\frac{1}{3}\tan^{-1}(3t - 4)$.

**3.** (A constant $C$ should be added in each case.)
(a) $\frac{1}{3}(x^2 - a^2)^{3/2}$. (b) $-(a^2 - x^2)^{1/2}$.
(c) $-(a^2 + x^2)^{-1/2}$. (d) $\frac{1}{2}(a^2 - x^2)^{-1}$.
(e) $\frac{1}{5}(x^2 - a^2)^{5/2}$. (f) $-\frac{1}{4}(x^2 + a^2)^{-2}$.

**5.** (A constant $C$ should be added in each case.)

(a) $\dfrac{-1}{1 - \cos x}$. (b) $\dfrac{-1}{3(2 + 3\sin x)}$. (c) $\tan^{-1}(\sin x)$.

(d) $\frac{2}{9}(1+x^3)^{3/2}$. (e) $\frac{1}{2}(1+x^4)^{1/2}$.

(f) $\frac{1}{3}\tan^3 x$. (g) $-\frac{1}{2}\cos x^2$. (h) $\frac{1}{12}\tan^{-1}\frac{3x}{4}$

(i) $\frac{1}{2}\tan^{-1} x^2$. (j) $\frac{1}{3}\sec^3 x$.

§ 5-5

1. (A constant $C$ should be added in each case.)

(a) $\frac{1}{3}\sin^{-1}\frac{3x}{4}$. (b) $\frac{1}{10}\tan^{-1}\frac{2x}{5}$.

(c) $\frac{1}{5}\sin^{-1}\frac{5x}{2\sqrt{2}}$. (d) $\frac{\sqrt{2}}{12}\tan^{-1}\frac{3x}{2\sqrt{2}}$.

(e) $\frac{\sqrt{3}}{3}\sin^{-1}\frac{\sqrt{3}x}{2}$. (f) $\frac{\sqrt{3}}{6}\tan^{-1}\frac{\sqrt{3}x}{2}$.

3. $4 - 2\sqrt{3} + \frac{2\pi}{3}$.

§ 5-6

1. (a) $31\frac{1}{4}$ ft. (b) $x = \frac{625 - v^2}{20}$.
   (c) $t = \frac{5}{2} - \frac{1}{10}\sqrt{625 - 20x}$.
3. (a) $\frac{88}{30}$ ft/sec². (b) $\frac{3}{16}$ miles. 5. (a) 4.92, 6.61, and 6.96 miles/sec, resp. (b) 4.68 and 6.80 miles/sec, resp.
7. (a) $v = 4(t-4)^2$. (b) $x = \frac{4}{3}t^3 - 16t^2 + 64t$.
   (c) $v = (512 - 6x)^{2/3}$. (d) $t = 4 - \frac{1}{2}(512 - 6x)^{1/3}$.
   (e) $t = 4$, $x = 85\frac{1}{3}$. (f) $v = 44.40$, $t = 0.67$.
9. (a) $v = \frac{1}{4}(10 - kx)^2$. (b) 1/50. (c) 200.
   (d) $x = \frac{500t}{t+20}$. (e) 500, 0.
11. (a) $k = 5$, $v^2 = 81 - 5x^2$. (b) $9/\sqrt{5}$, $2\pi/\sqrt{5}$.
   (c) $x = \frac{9}{\sqrt{5}}\sin\sqrt{5}t$.

§ 5-7

1. (a) The line $3x - 2y + 11 = 0$. (b) $x + y = 2$.
3. (a) All of $16y = x^2$. $y' = 3t/8$, $y'' = \frac{1}{8}$.
   (b) All of $4x = -y^2$. $y' = \frac{1}{t-2}$, $y'' = \frac{1}{2(t-2)^3}$.
   (c) All of $(y-2)^2 = 4(x+1)$, $y' = 2/t$, $y'' = -4/t^3$.

(d) Fourth quadrant part of $y^2 = x$. $y' = \dfrac{-1}{2t^2}$, $y'' = \dfrac{1}{4t^6}$.

(e) First and second quadrant parts of $y = 4 - x^2$. $y' = -4\cos\pi t$, $y'' = -2$.

(f) First quadrant part of $y = x^2$. $y' = 2\sqrt{1 + t^2}$, $y'' = 2$.

**9.** (a) $\left(\dfrac{v_0^2 \sin 2\alpha}{2g}, \dfrac{v_0^2 \sin^2 \alpha}{2g}\right)$. (c) $\alpha = 15°$, $t = 1.6$ sec (approx.), $\alpha = 75°$, $t = 6.0$ sec (approx). (d) Slant range is $\dfrac{v_0^2}{g(1 + \sin\theta)}$, for $\alpha = \dfrac{\pi}{4} + \dfrac{\theta}{2}$.

### § 5-8

**5.** (a) $x = a(\frac{1}{3})^{3/2}$, $y = a(\frac{2}{3})^{3/2}$. (b) $2a\omega/\sqrt{3}$.

**7.** $x = (a + b\cos\theta - b\cos\dfrac{a+b}{b}\theta$, $y = (a + b)\sin\theta - b\sin\dfrac{a+b}{b}\theta$.

## CHAPTER VI

### § 6-1

**1.** (a) 328, 168; (b) 240; (c) 284, 204; (d) 242. **3.** (a) 1600; (b) 3600; (c) 2450. **5.** (a) $\frac{501}{128}$, $\frac{397}{128}$. **7.** (a) $20\pi$, $12\pi$. (b) $16\pi$.

### § 6-4

**1.** (a) 60. (b) $2b^4/3$. (c) $-180.6$. (d) $\frac{14}{9}$. (e) $\frac{7}{36}$. (f) $\frac{1}{2}$. (g) $\frac{3}{100}$. (h) $\frac{52}{9}$. (i) $\frac{2}{5}$. (j) $2(\sqrt{3} - 1)$. (k) 0. (l) $7\pi/6$. (m) $\pi/6$. (n) $5\pi/72$. (o) $\dfrac{4\pi}{3} + \dfrac{\sqrt{3}}{2}$. (p) $\pi/8$.

**3.** (a) $\sin x$. (b) $\sqrt{y}$. (c) $\sqrt{1 + u^4}$. (d) $\tan^{-1} s$. (e) $-\cos x$. (f) $\dfrac{2}{1 + x^4}$.

**5.** (a) $\mu = 4$, $X = 2$. (b) $\mu = 16$, $X = 2\sqrt[3]{2}$.

(c) $\mu = \frac{4}{3}$, $X = \frac{16}{9}$. (d) $\mu = 2/\pi$, $X = \sin^{-1}\dfrac{2}{\pi}$.

(e) $\mu = \pi/4$, $X = \left(\dfrac{4 - \pi}{\pi}\right)^{1/2}$. (f) $\mu = \pi/6$, $X = \dfrac{2}{\pi}\sqrt{\pi^2 - 9}$.

### § 6-5

**1.** (a) $18\sqrt{2}$. (b) $\frac{243}{4}$. (c) 1. (d) 3. (e) 6. (f) 64. (g) $\frac{32}{15}\sqrt{2}$. (h) $\frac{64}{3}$. (i) 30. (j) 27. **3.** (a) 3. (b) $\frac{45}{5921}$. (c) 48. **5.** $10 - 5\sin^{-1}\dfrac{2}{\sqrt{5}}$. **7.** $a^2/6$. **9.** $4(3\sqrt{3} - \pi)$.

### § 6-7

**1.** $\frac{4000}{3}$. **3.** $32\pi/15$. **5.** (b) $\dfrac{\pi h^2}{3}(3a - h)$. (c) $4\pi ab^2/3$. (d) $4\pi a^2 b/3$. (e) $320\pi$.

(f) $16\pi$. (g) $2\pi/3$. (h) $64\pi$. (i) $\pi a^3/15$. **7.** (a) 256. (b) $64\sqrt{3}$. **9.** $\frac{abc}{24}(3\pi - 4)$.
**11.** $2\pi$. **13.** $16a^3/3$.

**15.** (a) $\frac{4}{3}(3\sqrt{3} + 2\pi)$. (b) $\frac{4}{3}(10 + 3\pi)\sqrt{2} - \frac{64}{3}$.

§ 6-8

**1.** (a) 27/4 ft. (b) 960 lb. **3.** (a) $-90kMm$ ergs. (b) 9 $kmM$ ergs. **5.** (a) $7\frac{1}{2}$ ft-lb. (b) 175/24 ft-lb. (c) 0. **7.** (a) $2\frac{1}{2}$ ft-lb. (b) 15/8 ft-lb.

§ 6-9

**1.** (a) $-7\sqrt{69}$ m/sec. (b) 490 joules. (c) 294 joules. (d) $-58.8$ m/sec., 6.1 m. (e) work = $+528.22$ joules. **3.** (a) 10,000 mile-pounds increase. (b) 7500 mile-pounds. (c) $V = mgx^2/2R$.

§ 6-10

**1.** (a) $Ma^2/3$. (b) $7Ma^2/3$. (c) $Ma^2/6$. (d) $\frac{1}{3}Ma^2 \sin^2 \theta$. **3.** Min. $I = Mb^2/18$, for $c = 2b/3$.

**5.** (a) $\int_a^b (x - c)^2 \sigma f(x)\,dx$. (b) $\frac{3}{7}MB^2$, $\frac{8}{35}MB^2$, $\frac{1}{5}MH^2$.

## *Review Problems, End of Chapter VI*

**1.** (a) Part of $x = 1 - 2y^2$ for which $-1 \le x \le 1$. $y' = -\frac{1}{4}\csc t$, $y'' = -\frac{1}{16}\csc^3 t$. (b) Part of $y = 2x^2 - 1$ for which $-1 \le y \le 1$. $y' = 4\cos t$, $y'' = 4$. (c) Part of $ay^2 = b^2(a - x)$ for which $x \ge 0$. $y' = -b/2a \cos t$, $y'' = -b/4a^2 \cos^3 t$. **3.** $-2p^2m^{-4}$. **5.** 14. **7.** Mean $= 4a^2/3$, max. $= 2a^2$. **9.** (a) $\pi b/4$. (b) $2H/3$.

**11.** (a) $V = a^2b \cos^{-1}\left(\frac{a - h}{a}\right) - b(a - h)\sqrt{2ah - h^2}$.

(b) $V = \frac{\pi h^2}{3}(3a - h)$.

# CHAPTER VII

§ 7-2

**1.** $90/\sqrt{89}$, $70/\sqrt{89}$. **3.** (a) $21x - 77y + 57 = 0$. (b) $(\frac{59}{33}, 0)$, $(0, \frac{59}{9})$. **5.** (a) $99x - 27y - 576 = 0$, $x - 3y - 4 = 0$, $2x + 3y - 14 = 0$. (b) $(6, \frac{2}{3})$. (c) 14/3. **7.** $\left(\frac{16}{5}, \frac{-11}{5}\right)$, $(-2, 3)$. **9.** A parabola. (a) $x^2 + y^2 - 2xy - 8x - 8y = 0$. (b) $(3, -1)$. (c) $(-1, 3)$. **11.** Four circles. $\left(\frac{-4}{11}, 0\right)$, $r = \frac{9}{11}$; $\left(\frac{-38}{7}, 0\right)$, $r = \frac{27}{7}$, $(1, -5)$, $r = 4$; $(1, \frac{135}{77})$, $r = \frac{108}{77}$.

§ 7-3

**1.** (a) $y = mx + 4$, $m = \frac{2}{3}$. (b) $2y = x + b$, $b = 4$. (c) $y = \frac{5}{2}x + b$, $b = \pm 2\sqrt{5}$. (d) $x \cos \alpha + y \sin \alpha = 5$, $-3x + 4y = 25$. (e) $y = m(x - 3)$, $m = -3$. **3.** (a) $57x - 38y = 38$. (b) $3x - 2y = 7$. (c) $3x + 19y + 26 = 0$. (d) $4x + 3y = 15$ and $3x - 4y = 15$. **7.** $\beta x + \alpha y = 2b^2$, where $(\alpha, \beta)$ is the point of tangency. **9.** $\alpha^{-1/3}x + \beta^{-1/3}y = b^{2/3}$, where $(\alpha, \beta)$ is the point of tangency.

§ 7-4

**1.** (a) 5. (b) $(1, 5)$, $(0, -2)$. **3.** (a) $y = -1$. (b) $x^2 + y^2 - 2(2 + k)y - 2k = 0$. (c) Radii are 2, $\sqrt{11}$, $\sqrt{5}/2$, $\sqrt{5}/2$. (d) $(0, -1 \pm \sqrt{5})$. **5.** $(\frac{29}{5}, \frac{8}{5})$. **7.** $11(x^2 + y^2) + 62x + 32y = 52$. **9.** (a) $7x + y = 3$. (b) Center $(-\frac{17}{6}, \frac{7}{6})$, $r = 13\sqrt{2}/6$.

§ 7-5

**1.** (a) Foci $(\pm 5, 0)$. (b) $c = 25$, $h = 144$, $k = -16$. $(\frac{39}{5}, \frac{48}{5})$. (c) Slope of ellipse is $-\frac{9}{13}$. **3.** (d) $x^2 = \dfrac{(c^2 + h)(c^2 + k)}{c^2}$, $y^2 = -\dfrac{hk}{c^2}$. (e) Slope of hyperbola is $\left[\dfrac{-k(c^2 + h)}{h(c^2 + k)}\right]^{1/2}$.

§ 7-6

**1.** (a) Ellipse. (b) Hyperbola. (c) Hyperbola. (d) Hyperbola. (e) Circle. **3.** (a) $u^2 + 5v^2 = 20$, ellipse. (b) $-u^2 + 11v^2 = 44$, hyperbola. (c) $u^2 - 3v^2 = 24$, hyperbola. **5.** $u^2 + 4v^2 = 4$, ellipse. **7.** $u^2 - v^2 = 32$. **9.** (a) $B = 0$, $A > 0$. (b) $A = 0$, $B \neq 0$. (c) $A^2 = B^2$, $A > 0$.

§ 7-7

**1.** (a) Ellipse, 12, $2\sqrt{6}$. (b) Hyperbola, $2\sqrt{7}$. (c) Two lines 4 units apart. (d) No locus. (e) One line. (f) Hyperbola, $\frac{8}{3}$. **3.** (a) Two lines, slope 3, 2 units apart. (b) Hyperbola, slope $-\frac{3}{4}$. (c) Ellipse, slope $\frac{3}{4}$. (d) Hyperbola, slope $-1/\sqrt{3}$. (e) Ellipse, slope $\frac{2}{3}$. (f) Hyperbola, slope $-5$.

§ 7-8

**1.** (a) $4V^2 - U^2 = 1$. (b) $V^2 = 4U$. (c) $4U^2 + V^2 = 16$. (d) $V = \pm 2$. (e) $U^2 + 4V^2 = 16$. (f) $y = 5x - 1$, $x + y + 5 = 0$. **3.** $9x^2 - 42xy + 49y^2 - 72x - 24y + 144 = 0$. Slope of axis is $\frac{3}{7}$. Second equation describes line through $(3, 3)$ and $(4, 0)$.

## CHAPTER VIII

§ 8-1

**1.** (a) 5. (b) $-6$. (c) $-4$. (d) $-\frac{1}{2}$. (e) $-4$. (f) $\frac{3}{2}$.

§ 8-5

**1.** (a) $\dfrac{10x}{x^2+9}$.

(b) $\dfrac{2(a-x)}{2ax-x^2}$.

(c) $\operatorname{ctn} x$.

(d) $6 \csc x$.

(e) $x + 2x \log x$.

(f) $\dfrac{2x^2}{x^2-16} + \log(16-x^2)$.

(g) $\dfrac{4 \log x^2}{x}$.

(h) $\dfrac{2 \log x}{x^2} - \left(\dfrac{\log x}{x}\right)^2$.

(i) $2 \tan 2x$.

(j) $4 \sec 4x$.

**3.** (a) $(x \log x)^{-1}$.

(b) $-x^{-1} \sin(\log x)$.

(c) $3x^2 \log x$.

(d) $2 \sin(\log x)$.

(e) $a^2xe^{ax}$.

(f) $(1+e^x)^{-1}$.

(g) $2(e^{2x}-e^{-2x})^{-1}$.

(h) $2(e^{2x}+e^{-2x})^{-1}$.

(i) $\dfrac{-e^{-x/2}}{2\sqrt{1-e^{-x}}}$.

(j) $(1+e^{-x})^{-1/2}$.

**9.** Flux decreases as $r$ increases. Concave upward if $0 < r < h/2$.

§ 8-6

**1.** $e^{-(t/5)\log(5/2)}$. **3.** (a) $i = i_0e^{-Rt/L}$. (b) 5.18 sec. **5.** (a) 3450 (approx.) (b) 21.6.

**9.** (a) $p = 1600e^{-(t/2)\log 2}$ rpm. (b) 8 min. $\dfrac{6400\pi}{\log 2}(1 - e^{-(t/2)\log 2})$ radians.

**11.** 180 lb. **13.** $6\frac{1}{4}\%$. **15.** About 6.48%.

**17.**
$$V = \frac{cV_0}{r_0} + \left(1 - \frac{c}{r_0}\right)V_0e^{-r_0t/V_0}.$$

## CHAPTER IX

§ 9-2

**1.** (a) $x = \pm 1$.

**5.** (a) $\dfrac{2}{\sqrt{4x^2-4x+2}}$.

(b) $\dfrac{3}{\sqrt{9x^2+30x+24}}$.

(c) $\dfrac{5}{25x^2-20x+3}$.

(d) $-\operatorname{csch} x$.

(e) $\sec x$ if $0 < x < \pi/2$.

(f) $\sec x$ if $|x| < \pi/2$.

§ 9-3

1. (a) $\frac{1}{6}\sinh^3 2x + C.$ (b) $-\frac{1}{3}\operatorname{sech}^3 x + C.$
(c) $\frac{1}{3}\cosh^3 x - \cosh x + C.$ (d) $\frac{1}{4}\sinh 4x + \frac{1}{6}\sinh^3 4x + \frac{1}{20}\sinh^5 4x + C.$

3. (a) $\dfrac{-x}{a^2\sqrt{x^2 - a^2}} + C.$ (b) $\dfrac{\sqrt{x^2 - a^2}}{a^2 x} + C.$

9. (a) Symmetry with respect to $y$-axis. $y = 0$ is asymptote. (c) Limit is $\pi/2$.

## *Review Problems, End of Chapter IX*

1. Hyperbola, asymptotes $x = 1, y = 2$. 3. Hyperbola $(x - 4)(y - 2) = 8$, center $(4, 2)$. 7. $x^2 + 2xy = 16$. Asymptotes $x = 0$, $x + 2y = 0$. 9. $x - 2y + 4 = 0$. 11. $7x - y = 7$, $y = x + 5$. 13. (a) $2py = 2x_0x - x_0^2$. (b) $x_1 = -p^2/x_0$. Intersection $\left(\dfrac{x_0^2 - p^2}{2x_0}, \dfrac{-p}{2}\right)$. 17. Max. at $x = \pm 1$, min. at $x = 0$. $2x^4 - 5x^2 + 1 = 0$. 19. $eV/R$. 21. $-6.3$ lb/ft²/sec.

# CHAPTER X

(Add $C$ to all indefinite integral answers.)

§ 10-2

1. (a) $-\frac{1}{3}\log|\cos(3x - 4)|.$ (d) $-\log(e^{-x} + 2).$
(b) $-e^{-\sin x}.$ (e) $\tan^{-1} e^x.$
(c) $\log(\log x).$ (f) $-\frac{1}{4}\sin^{-1}\left(\dfrac{\cos^2 2x}{3}\right).$

§ 10-3

1. $\frac{1}{2}\sin^{-1}\dfrac{8x - 9}{9}.$ 3. $-\sqrt{8 - 2x - x^2}.$

5. $\dfrac{1}{\sqrt{5}}\tan^{-1}\dfrac{3x + 7}{\sqrt{5}}.$ 7. $-\sqrt{6x - x^2 - 5} + \sin^{-1}\dfrac{x - 3}{2}.$

9. $-\sqrt{4x - x^2} + 2\sin^{-1}\dfrac{x - 2}{2}.$

11. $\dfrac{x}{50(4x^2 + 25)} + \dfrac{1}{500}\tan^{-1}\dfrac{2x}{5}.$

13. $\dfrac{x - 2}{3(x^2 - x + 1)} + \dfrac{2}{3\sqrt{3}}\tan^{-1}\dfrac{2x - 1}{\sqrt{3}}.$

15. $\dfrac{(x + 2)[12x^2 + 48x + 173]}{5000(4x^2 + 16x + 41)^2} + \dfrac{3}{50{,}000}\tan^{-1}\dfrac{2x + 4}{5}.$

## § 10-4

1. $\dfrac{x^2}{2} - 8x + 68 \log |x + 8|.$

3. $\dfrac{x^2}{2} - 4x + \dfrac{1}{2} \log |x| + \dfrac{31}{2} \log |x + 4|.$

5. $\dfrac{x^2}{2} + x + \dfrac{1}{5} \log |(x - 1)(x^2 + 4)^2| + \dfrac{2}{5} \tan^{-1} \dfrac{x}{2}.$

7. $\log |x - 2| - \dfrac{2(2x - 3)}{(x - 2)^2}.$

9. $\dfrac{x}{2} - \dfrac{1}{5} \log |x| + \dfrac{29}{20} \log |2x - 5|.$

11. $\dfrac{1}{11} \log \left|\dfrac{5 - x}{6 + x}\right|.$

13. $\dfrac{1}{8} \log \left|\dfrac{(x - 1)^3}{(x + 1)(x^2 + 1)}\right| - \tan^{-1} x + \dfrac{1 - 2x}{4(x^2 + 1)}.$

15. $\dfrac{1}{10} \log \dfrac{(x - 1)^2}{x^2 + x + 3} - \dfrac{3}{5\sqrt{11}} \tan^{-1} \dfrac{2x + 1}{\sqrt{11}}.$

17. $\dfrac{3}{2} \log (85 + 60\sqrt{2}).$

## § 10-5

1. (a) $(x - 1)e^x.$
   (b) $-(x^2 + 2x + 2)e^{-x}.$
   (c) $\frac{1}{8}(4x^3 - 6x^2 + 6x - 3)e^{2x}.$
   (d) $\frac{1}{2}x \sin 2x + \frac{1}{4} \cos 2x.$
   (e) $(2 - x^2) \cos x + 2x \sin x.$
   (f) $(x^3 - 6x) \sin x + 3(x^2 - 2) \cos x.$

5. (a) $$\int \frac{(\log x)^n}{x}\, dx = \frac{(\log x)^{n+1}}{n + 1} \quad \text{if } n \neq -1,$$
   $$= \log |\log x| \quad \text{if } n = -1.$$
   (b) $\dfrac{x^{11}}{(11)^5} [(11 \log x)^4 - 4(11 \log x)^3 + 12(11 \log x)^2 - 24(11 \log x) + 24].$

## § 10-6

5. (a) $\tan x - x.$
   (b) $\frac{1}{2} \tan^2 x + \log |\cos x|.$
   (c) $\frac{1}{3} \tan^3 x - \tan x + x.$
   (d) $\frac{1}{3} \tan^3 x + \tan x.$

(e) $\frac{1}{5}\tan^5 x + \frac{2}{3}\tan^3 x + \tan x$.
(f) $\frac{1}{5}\tan^5 x + \frac{1}{3}\tan^3 x$.

**7.** (a) $\frac{1}{3}\sec^3 x - \sec x$.
(b) $\frac{1}{7}\sec^7 x - \frac{2}{5}\sec^5 x + \frac{1}{3}\sec^3 x$.

**9.** (b) $\int \tan^6 2x\,dx = \frac{1}{8}\tan^4 2x - \frac{1}{4}\tan^2 2x + \frac{1}{2}\log|\sec 2x|$.

$$\int \text{ctn}^6 3x\,dx = -\tfrac{1}{15}\text{ctn}^5 3x + \tfrac{1}{9}\text{ctn}^3 3x - \tfrac{1}{3}\text{ctn}\,3x - x.$$

**11.** (a) $\dfrac{\cos x \sin^5 x}{6} - \dfrac{\cos x \sin^3 x}{24} - \dfrac{\cos x \sin x}{16} + \dfrac{x}{16}$.

(b) $-\dfrac{\sin x \cos^7 x}{8} + \dfrac{\sin x \cos^5 x}{48} + \dfrac{5 \sin x \cos^3 x}{192} + \dfrac{5 \sin x \cos x}{128} + \dfrac{5x}{128}$.

(c) $-\frac{1}{3}\text{ctn}^3 x$.

(d) $\dfrac{\sin^3 x}{\cos x} + \frac{3}{2}\sin x \cos x - \frac{3}{2}x$.

**13.** (a) $-\frac{1}{2}\text{ctn}\,\theta \csc\theta + \frac{1}{2}\log|\csc\theta - \text{ctn}\,\theta|$.
(b) $\log|\tan x|$ or $\log|\csc 2x - \text{ctn}\,2x|$.
(c) $\frac{1}{6}\tan^3 2x + \frac{1}{10}\tan^5 2x$.
(d) $-\frac{1}{3}(\frac{1}{5}\text{ctn}^5 3x + \frac{2}{7}\text{ctn}^7 3x + \frac{1}{9}\text{ctn}^9 3x)$.
(e) $-\frac{1}{4}\csc^4 x$.

(f) $\dfrac{\cos^5 x}{2\sin^3 x} - \frac{5}{6}\text{ctn}^3 x + \frac{5}{2}\text{ctn}\,x + \frac{5}{2}x$.

**15.** (a) $-\frac{1}{6}\cos 3x + \frac{1}{2}\cos x$.
(b) $-\frac{1}{14}\cos 7x + \frac{1}{2}\cos x$.
(c) $-\frac{1}{24}\sin 6x + \frac{1}{8}\sin 4x - \frac{1}{8}\sin 2x$.
(d) $-\frac{1}{28}\cos 7x - \frac{1}{6}\cos 3x + \frac{1}{4}\cos x$.

**17.** (b) $\dfrac{8x^4 - 3}{32}\sin^{-1} x + \dfrac{2x^3 + 3x}{32}\sqrt{1 - x^2}$.

## § 10-7

**5.** (a) $\dfrac{x - a}{a^2\sqrt{2ax - x^2}}$.

(b) $\dfrac{x - 3a}{2}\sqrt{x^2 + 2ax} + \dfrac{3a^2}{2}\log(x + a + \sqrt{x^2 + 2ax})$.

**7.** $ab\left[\dfrac{\pi}{4} - \dfrac{1}{2}\sin^{-1}\dfrac{3}{5} - \dfrac{2}{5}\right]$.

## § 10-8

**1.** (a) $\frac{1}{12}(2x - 3)\sqrt{3 + 4x}$.

(b) $\dfrac{x^2 - 2a^2}{\sqrt{x^2 - a^2}}$.

(c) $\frac{1}{2}(x + 3)(3x - 5)^{1/3}$.

(d) $\dfrac{2a^2 + 6 - x^2}{3}\sqrt{a^2 + x^2}$.

(e) $\frac{2}{15}(3x^2 - 8x + 32)\sqrt{x + 2}$.

(f) $\frac{3}{2} \log |x^{2/3} - 1|$.

(g) $\log \dfrac{\sqrt{1 + 4x} - 1}{\sqrt{1 + 4x} + 1}$.

(h) $\dfrac{1}{a} \log \dfrac{|x|}{a + \sqrt{x^2 + a^2}}$.

(i) $-\dfrac{\sqrt{a^2 - x^2}}{2a^2x^2} + \dfrac{1}{2a^3} \log \dfrac{a - \sqrt{a^2 - x^2}}{|x|}$.

(j) $\dfrac{2}{\sqrt{5}} \tan^{-1}\left[\dfrac{\tan (x/2)}{\sqrt{5}}\right]$.

(k) $\sqrt{2} \tan^{-1}\left(\dfrac{1 + 3 \tan (x/2)}{\sqrt{2}}\right)$.

(l) $x - \dfrac{8}{\sqrt{15}} \tan^{-1}\left(\dfrac{4 \tan (x/2) + 1}{\sqrt{15}}\right)$.

**§ 10-9**

**1.** (a) $\dfrac{\sqrt{2}}{2}\left(\tan^{-1} \dfrac{3\sqrt{2}}{4} - \tan^{-1} \dfrac{\sqrt{2}}{4}\right)$.

(b) $\dfrac{\pi + 2}{512}$. (c) $367\frac{11}{56}$. (d) $\dfrac{4^9}{35}$. (e) $\dfrac{\pi a^6}{32}$. (f) $15\pi$.

## CHAPTER XI

**§ 11-1**

**1.** (a) $\frac{8}{27}(10\sqrt{10} - 1)$. (b) $\dfrac{\sqrt{5}}{2} + \log\left(\dfrac{3 + \sqrt{5}}{2}\right)$. (c) $\log (2 + \sqrt{3})$. (d) $-\frac{3}{4} + \log 7$. (e) $2 \sinh 1$. (f) $\frac{2}{27}(10\sqrt{10} - 1)$. (g) $\frac{1}{2}[\sqrt{6} + \log (\sqrt{3} + \sqrt{2})]$. (h) $\frac{9}{8}$. **3.** (a) $\frac{1}{27}(104\sqrt{13} - 125)$. (b) $4 + \frac{1}{2} \log \frac{3}{2}$. (c) $\dfrac{25\pi}{6}$. (d) $4[\sqrt{2} + \log (\sqrt{2} + 1)]$. (e) $\sqrt{2}(e^2 - 1)$. (f) $4\pi$. (g) $3\sqrt{2} + \frac{1}{2} \log (3 + 2\sqrt{2})$. (h) $a \log 2$. **5.** The same integrals with $\sin \theta$ and $\cos \theta$ exchanged.

§ 11-2

**3.** $4\pi a^2b/3$. **5.** (a) $576\pi$. (b) $16\pi \log 2$. (c) $\pi$. (d) $\pi(1 - e^{-1/2})$. (e) $2\pi$. (f) $2\pi^2$. (g) $144\pi\sqrt{3}$. (h) $64\pi$. (i) $5\pi/6$. (j) $10\pi/3$. (k) $27\pi$.

§ 11-4

**1.** (a) $208\pi/3$. (b) $\frac{\pi}{9}(17\sqrt{17} - 1)$. (c) $104\pi/3$. (d) $2\pi[\sqrt{2} + \log(\sqrt{2} + 1)]$. (e) $\pi[2 + \sinh 2]$. **3.** $\pi m\sqrt{1 + m^2}$. **5.** $2\pi[\sqrt{2} + \log(1 + \sqrt{2})]$. **7.** (a) $2\pi b^2 + \dfrac{2\pi ab}{e}\sin^{-1} e$. (b) $2\pi a^2 + \dfrac{2\pi b^2}{e}\tanh^{-1} e$.

§ 11-5

**1.** $(\frac{9}{8}, \frac{5}{6})$.

§ 11-6

**1.** $\bar{x} = 3h/4$. **3.** $\bar{x} = \dfrac{3}{4}\dfrac{(2a - h)^2}{3a - h}$. **5.** $\bar{y} = 3b/8$. **7.** $\bar{y} = \frac{4}{3}$. **9.** $\bar{y} = \frac{8}{9}$. **11.** (a) $\bar{y} = \frac{27}{10}$. (b) $\bar{x} = \frac{27}{16}$. (c) $\bar{x} = \frac{39}{20}$. (d) $\bar{y} = \frac{21}{10}$. **13.** (a) $\bar{x} = 27a/16$. (b) $\bar{y} = 3\sqrt{3}a/8$. (c) $\bar{y} = \dfrac{5a}{24[\sqrt{3} - \log(2 + \sqrt{3})]}$.

§ 11-7

**1.** $\bar{x} = \frac{3}{5}$, $\bar{y} = \frac{3}{4}$. **3.** $\bar{x} = \frac{6}{5}$, $\bar{y} = 0$. **5.** $\bar{x} = 4a/3\pi$. **7.** The intersection of the medians. **9.** (a) $(\frac{5}{2}, 5)$. (b) $(\frac{12}{5}, \frac{3}{2})$. (c) $(\frac{1}{2}, \frac{8}{5})$. (d) $(\frac{16}{15}, -\frac{24}{35})$. (e) $(\frac{88}{39}, \frac{50}{39})$. (f) $(\frac{34}{7}, \frac{37}{12})$. **11.** $\bar{x} = \bar{y} = 256a/315\pi$.

§ 11-8

**1.** (a) $128w/5$. (b) $544w/15$. (c) $648w/5$. (d) $9\sqrt{3}w$. (e) $6\sqrt{3}w$. (f) $28w$; $16w$. (g) $39w$; $51w$. (h) $14w/3$. (i) $\dfrac{4w}{3}(3\pi - 4)$; $\dfrac{4w}{3}(3\pi + 4)$. **3.** $wh^2b/2$. **5.** Almost $5\pi$ tons. **7.** $16w/15$. **9.** $12\pi w$. **11.** $\frac{2}{3}$ the way down the gate.

§ 11-9

**1.** $\bar{x} = a/2$. **3.** $\bar{x} = \dfrac{4(8 - 3\sqrt{3})}{2\pi + 3\sqrt{3}}$. **5.** $\bar{y} = 2a/\pi$.

**7.** $\bar{y} = \dfrac{18\sqrt{5} - \log(2 + \sqrt{5})}{32\sqrt{5} + 16\log(2 + \sqrt{5})} \sim 0.4$.

**9.** $(\pi a/4, a/2)$.

## CHAPTER XII

§ 12-1

9. (a) $\pm 3\sqrt{3}/2$. (b) $\pm 1$.

§ 12-2.

1. (a) $p = a(1 - e^2)/e$. 3. (2, 0) and (−2, 0). 5. (a) and (d) ellipses; (b) and (c) parabolas; (e) and (f) hyperbolas. 9. Either 15 or 45 million miles.

§ 12-3

1. (a) $b^2\, d\theta^2$. (b) $2a^2(1 - \cos\theta)\, d\theta^2$.
 (c) $\dfrac{2a^2\, d\theta^2}{(1 + \cos\theta)^3}$. (d) $a^2 \csc 2\theta\, d\theta^2$.
 (e) $4 \sin^2 \dfrac{\theta}{2}\, d\theta^2$. (f) $16 \sin^4 \dfrac{\theta}{3}\, d\theta^2$.
3. 16.27.
5. (a) $\frac{2}{3}(13\sqrt{13} - 8)$.
 (b) $\dfrac{1}{2}\left[2\sqrt{5} - \sqrt{2} + \log \dfrac{2 + \sqrt{5}}{1 + \sqrt{2}}\right]$.
 (c) $2\left[\sqrt{5} - \dfrac{\sqrt{17}}{4} + \log \dfrac{8 + 2\sqrt{17}}{1 + \sqrt{5}}\right]$.
 (d) $\frac{2}{3}(3 - \sqrt{3})$.
 (e) $5\pi/3$. (f) $4\sqrt{2}(e^{\pi} - 1)$.
7. $32\pi a^2/5$. 9. $\left(\dfrac{4a}{5}, 0\right)$. 13. (a) At right angles at two points. (b) At right angles at origin; acute angle between curves is $\pi/3$ at $\theta = \pi/6$ and $\theta = 5\pi/6$. (c) Acute angle between curves is $\pi/4$ at origin, $\pi/3$ at $\theta = \pi/6$ and $\theta = 5\pi/6$.

§ 12-4

1. (a) 16. (b) $\pi a^2$. (c) $3\pi a^2/2$. (d) $24\pi a^2$. (e) $\pi a^2/2$. (f) $16\pi$. (g) $6\pi$. (h) $\dfrac{\pi}{2}(2a^2 + b^2)$.
 3. (a) $16\pi - 24\sqrt{3}$, $32\pi + 24\sqrt{3}$. (b) $\pi - \frac{3}{2}\sqrt{3}$, $\pi + 3\sqrt{3}$. (c) $\frac{1}{2}(\pi - 3)$, $\frac{3}{2}(\pi + 1)$. (d) $\frac{1}{6}(5\pi - 9\sqrt{3})$, $\frac{1}{6}(25\pi + 9\sqrt{3})$. 5. $2p^2/3$. 7. $\dfrac{a^2}{4}(5\pi - 8)$.

### *Review Problems, End of Chapter XII*

1. (a) $10\pi/3$. (b) $5\pi$. (c) $\frac{3}{2} - 2\sin^{-1}\frac{4}{5}$. (d) $\frac{16}{3}$. (e) $5\sqrt{15} - 5\log(4 + \sqrt{15})$.
 3. $\pi^2/2$. 5. $5\pi^2 a^3$. 7. $\bar{x} = \dfrac{2p}{35}(5 + 3\sqrt{2})$. 9. $\dfrac{1 + \sqrt{2}}{3}a$. 13. (a) $\theta = \pi/12$ and $\pi/6$. (c) $\theta = \pi/12$ and $\pi/6$.

## CHAPTER XIII

### § 13-2

**1.** (a) $5\mathbf{i} - 12\mathbf{j}$, length 13. (b) $-24\mathbf{i} + 18\mathbf{j}$, length 30. **3.** 75°, −15°. **5.** $\frac{1}{\sqrt{5}}(\mathbf{i} - 2\mathbf{j})$. **9.** Line through tip of **A** parallel to **B**. **15.** $25x^2 + 9y^2 = 225$. Counterclockwise. $d\mathbf{R}/dt = -\frac{48}{5}\mathbf{i} + 12\mathbf{j}$, $d^2\mathbf{R}/dt^2 = -16\mathbf{R}$.

### § 13-3

**1.** (a) Parabola $(y + 2)^2 = x - 4$. $\mathbf{V} = 4\mathbf{i} + \mathbf{j}$ at $t = 2$ (crossing of $x$-axis). Speed least at $t = 0$. (b) Circle $x^2 + (y - a)^2 = a^2$. Point goes counterclockwise at constant speed. $\mathbf{V} = -\pi a\mathbf{i} + \pi a\sqrt{3}\,\mathbf{j}$ at $t = \frac{1}{3}$. (c) $y^2 = x^3$; cusp at $t = 0$. $\mathbf{V} = 4\mathbf{i} + 12\mathbf{j}$ at $t = 2$. (d) Ellipse $16(x - 5)^2 + 25y^2 = 400$. Clockwise motion. Period is 2 time units. Max. speed $5\pi$, min. speed $4\pi$. $\mathbf{V} = -\frac{5\pi}{2}\mathbf{i} - 2\pi\sqrt{3}\mathbf{j}$ at $t = \frac{2}{3}$. (e) $xy = 1$ (hyperbola). Speed least at $t = 0$. (f) $y = 8\sqrt{2}\log x$. Speed least at $t = 2$. (g) Parabola $4x^2 - 4xy + y^2 = 625x$. Speed least at vertex, where $t = 5$. Axis of parabola is $y = 2x - 175$. **3.** (a) $-3(\mathbf{i} + 2\sqrt{2}\mathbf{j})$. (b) $5\sqrt{3}$ units/sec. **7.** $Q$ moves upward $\frac{\sqrt{x+1}}{x}$ units/min. **9.** Vertical component $16\sqrt{2}\sin\theta\sqrt{\cos\theta}$; horizontal component $-16\sqrt{2}(\cos\theta)^{3/2}$.

### § 13-4

**1.** (a) $A_x = 0$, $A_y = -32$, $A_T = \dfrac{8(t-2)}{[1 + 64(t-2)^2]^{1/2}}$.

(b) $A_x = 2$, $A_y = 0$, $A_T = \dfrac{4t}{\sqrt{4t^2 + 1}}$.

(c) $A_x = 2$, $A_y = 6t$, $A_T = \dfrac{-2(2 + 9t^2)}{\sqrt{4 + 9t^2}}$ if $t < 0$.

(d) $A_x = 3\pi^2/\sqrt{2}$, $A_y = 5\pi^2/\sqrt{2}$, $A_T = -8\pi^2/\sqrt{17}$.

(e) $A_x = -4a\cos 2t$, $A_y = -4a\sin 2t$, $A_T = 0$.

(f) $A_x = \dfrac{4t(t^2 - 3)}{(1 + t^2)^3}$, $A_y = \dfrac{4(3t^2 - 1)}{(1 + t^2)^3}$, $A_T = \dfrac{-4t}{(1 + t^2)^2}$.

**3.** $x = 40\sqrt{2}t$, $y = 40\sqrt{2}t - 16t^2$. $\mathbf{A} = -32j$. $A_T = -16\sqrt{2}$ when $t = 0$.

**5.** $\mathbf{V} = \frac{\pi a}{20}(\mathbf{i}\cos\theta - \mathbf{j}\sin\theta)$, $\mathbf{A} = -\frac{\pi^2 a}{400}(\mathbf{i}\sin\theta + \mathbf{j}\cos\theta)$. Locus is a circle of radius $a$; **A** points toward center of circle and is constant in length.

§ 13-5

**1.** (a) $\frac{192}{125}$. Min. at $x = (\frac{1}{45})^{1/4}$.

(b) $\frac{12}{(17)^{3/2}}$. Min. at $x = (\frac{1}{84})^{1/6}$.

(c) $\frac{-4\sqrt{15}}{25}$. Min. at $x = \pi/2$.

(d) $-\sqrt{2}/2$. Min. at $x = 0$.

(e) $\frac{1}{8}$. Min. at $t = 1$.

(f) $\frac{16}{7\sqrt{7}}$. Min. at (1, 1) and (1, 5).

(g) $\frac{-80}{79\sqrt{79}}$. Min. at $t = \pm\frac{\pi}{2}$.

**3.** (a) $e^{\theta}\sqrt{2}$. (b) $a|\theta|$. (c) $y^2/a$. (d) $4a\left|\sin\frac{\theta}{2}\right|$. **5.** 1 radian/sec. **11.** (a) $\sqrt{6}/3$. (b) $\sqrt{2}$. (c) $\sqrt{2}$. (d) $v^2\sqrt{2}/8$. (e) $\sqrt{2}/3$. **13.** (a) $375\sqrt{3}$ lb, 125 lb. (b) $250\sqrt{2}$ lb.

§ 13-6

**1.** (a) $V_r = V_\theta = 15\sqrt{2}$; $A_\theta = -A_r = 90\sqrt{2}$.

(b) $V_r = V_\theta = 2\pi$; $A_r = 0$, $A_\theta = 8\pi^2$.

(c) $V_r = -\pi a \sin\theta$, $V_\theta = \pi a(1 + \cos\theta)$; $A_r = -\pi^2 a(1 + 2\cos\theta)$, $A_\theta = -2\pi^2 a \sin\theta$.

(d) $V_r = -a\sin\frac{\theta}{2}$, $V_\theta = a\cos\frac{\theta}{2}$; $A_r = -\frac{3a}{4}$, $A_\theta = -\frac{3a}{4}\tan\frac{\theta}{2}$.

(e) $V_r = 4\pi\cos 2\theta$, $V_\theta = 2\pi(2 + \sin 2\theta)$; $A_r = -4\pi^2(2 + 5\sin 2\theta)$, $A_\theta = 16\pi^2\cos 2\theta$.

§ 13-7

**1.** (a) $(-2, 3)$. (b) $(4, 4)$. (c) $(a, -3a/2)$.

**3.** $X = -x(1 + \frac{9}{2}x)$, $Y = \frac{4}{3}\sqrt{x}(1 + 3x)$.

## CHAPTER XIV

§ 14-3

**1.** $A_n$ increases; least upper bound is $\frac{16}{3}$. $S_n$ decreases; greatest lower bound is $\frac{16}{3}$. **3.** (a) $x_{n+1} < x_n$ if $n \geq 4$. (b) $x_{n+1} < x_n$ if $n \geq 3$. (c) $x_{n+1} < x_n$ if $n > 4$. (d) $x_n < x_{n+1}$ if $n > 4$. (e) $x_{n+1} > x_n$ if $n \geq 100$. (f) $x_{n+1} < x_n$ if $n \geq 6$.

### § 14-4

**1.** (a) 0. (b) 0. (c) 0. (d) $\frac{5}{3}$. (e) $10^4$. (f) $1/\sqrt{2}$.

### § 14-5

**1.** (a) 1. (b) $\frac{2}{3}$. (c) 0. (d) 15. (e) 3. (f) $\frac{8}{3}$. **7.** (a) $e^{-1}$. (b) 1. (c) 1. (d) $-1$. (e) $e^2$. (f) $e^{-2}$. **9.** Limit $-\infty$ as $x \to 0^-$, 0 as $x \to 0^+$.

## CHAPTER XV

### § 15-1

**1.** (a) $\frac{4}{9}$. (b) $\frac{132}{999}$. (c) $\frac{173}{55}$. (d) 3. **3.** (a) Div. (b) Conv. (c) Div. (d) Conv. (e) Div. (f) Div. **5.** Convergent, with sum $A + B$.

### § 15-2

**1.** (b) $10^4$ terms.

### § 15-3

**7.**
$$\sin x = \sin a + \cos a \frac{x-a}{1!} - \sin a \frac{(x-a)^2}{2!} - \cos a \frac{(x-a)^3}{3!} + \sin a \frac{(x-a)^4}{4!} + \cdots,$$

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

### § 15-4

**1.** (a) $1 - 2x + 3x^2 - 4x^3 + \cdots$.

(b) $\frac{2\cdot 1}{2} + \frac{3\cdot 2}{2} x + \frac{4\cdot 3}{2} x^2 + \cdots$.

(c) $1 - \frac{3}{2} x + \frac{3\cdot 5}{2\cdot 4} x^2 - \frac{3\cdot 5\cdot 7}{2\cdot 4\cdot 6} x^3 + \cdots$.

(d) $1 + \frac{3}{2} x + \frac{3}{8} x^2 - \frac{3}{8}\frac{1}{6} x^3 + \frac{3}{8}\frac{1\cdot 3}{6\cdot 8} x^4 - \frac{3}{8}\frac{1\cdot 3\cdot 5}{6\cdot 8\cdot 10} x^5 + \cdots$.

**5.** (a) $\sqrt{x} = 3 + \frac{x-9}{6} - \frac{(x-9)^2}{216} + \frac{(x-9)^3}{16X^{5/2}}$, $X$ between 9 and $x$. (b) $|R_2(x)| < \frac{1}{3888}$. (c) 3.162.

### § 15-6

**3.** (a) Conv. (b) Div. (c) Conv. (d) Conv. (e) Div. (f) Conv.

§ 15-7

1. (a) Conv. (b) Conv. if $p > 1$, div. if $p \leq 1$. (c) Div. (d) Div. (e) Div. (f) Conv. 5. Between 0.009001 and 0.009101.

§ 15-8

1. Theorem 15-G applicable and series convergent in (b), (c), (d), (e), (g), (h). Also applicable in (f) if $n \geq 3$. Series (a) div. ($a_n$ does not $\to 0$). 3. 0.905 with error less than $\frac{5}{8}(10^{-4})$.

§ 15-9

1. (a) Abs. conv. if $|x| < 1$, div. if $|x| > 1$. (b) $x = 1$: conv. if $p > 1$, div. if $p \leq 1$. $x = -1$: cond. conv. if $p > 0$, div. if $p \leq 0$. 3. (a) $|x| < 3$. (b) All values of $x$. (c) $-1 \leq x \leq 1$. (d) $-5 < x < 3$. (e) $-2 < x \leq 2$. (f) $-1 \leq x < 1$. (g) All values of $x$. (h) $|x| > 0$. 5. 0.570, using four terms.

§ 15-10

1. $x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7} + \cdots$.

15. $1 - 2x + x^2 + 2x^3 - 4x^4 + 2x^5 + 3x^6 + \cdots$.
17. $x - \frac{1}{2}x^2 + \frac{1}{6}x^3 - \frac{1}{12}x^4 + \frac{1}{24}x^5 - \cdots$.

## *Review Problems, End of Chapter XV*

3. At $x^2 = \dfrac{6a^2 + \sqrt{81a^4 + 5b^4}}{15}$, which exceeds $a^2$. No. 5. Monotonic, but not bounded. 9. $w \to PR^2/16\pi D$ and $dw/dr \to 0$ as $r \to 0$. $r/R = 1/e$ at inflection. 13. $R_2(x) = 161{,}700(1 + X)^{97}x^3$, $X$ between 0 and $x$. $|R_2| < 1617(10^{-7})$. $(0.999)^{100} = 0.905$ to three decimal places.

17. $p\left[\frac{1}{2} + \frac{1}{4}\left(\frac{h}{p}\right) - \frac{1}{8}\left(\frac{h}{p}\right)^2 + \frac{1}{16}\left(\frac{h}{p}\right)^3 - \cdots\right]$.

# CHAPTER XVI

§ 16-1

1. (a) 6.08. (b) 7.06. (c) 9.92. (d) 4.95. (e) 1.975. (f) $32(10^6)$. 3. 1.6394. Too large; tangent above curve. 5. (a) 216 cu in. (b) 36 sq in. (c) 0.52%, 1.04%, 1.56%.

§ 16-2

1. $x = 0.739$. 3. First method: $x = 1.28$, $x_2 = 1.27$. 5. $x = 0.64$.

### § 16-3

**1.** No. Converges to root between $\frac{1}{2}$ and 1. **3.** $-0.88$ and 1.35. **5.** 1.3054. **7.** 1.59. **9.** $\mu = 0.40$.

### § 16-4

**1.** (a) 0.77. (b) 0.76. **3.** $0.91\pi$. **9.** 1.852.

## CHAPTER XVII

### § 17-1

**1.** (a) 10. (b) $-2$. (c) 16.
**3.** (a) 2nd row = 3(1st row); 1st col. = $-2$(2nd col.).
(b) 2nd row = $\frac{3}{2}$(1st row); 1st col. = $\frac{2}{3}$(2nd col.).
(c) 2nd row = 0(1st row); 2nd col. = $\frac{4}{3}$(1st col.).
(d) 2nd row = $-\frac{3}{2}$(1st row); 2nd col. = 0(1st col.).
**5.** Sign changes when rows are exchanged; likewise for exchange of columns. **7.** (a) $-8$. (b) 6.

### § 17-2

**1.** (a) $-35$. (b) 186. (c) 0. (d) $-3$. (e) 0. (f) 29. (g) $-9$. (h) 0.

### § 17-3

**7.** (a) $-480$. (b) $-38$. (c) 9. (d) 29.

### § 17-4

**1.** (a) $(1, \frac{1}{2}, \frac{1}{3})$. (b) $(2, -2, 3)$. (c) $(3, 12, -6)$. (d) $(0, 4, -5)$.
**3.** (a) 4(1st col.) − (2nd col.) + 5(3rd col.) = col. of zeros.
(b) 2(1st row) + (2nd row) − (3rd row) = row of zeros.

## CHAPTER XVIII

### § 18-1

**1.** (a) $x = -1, 4; y = -2, 3; z = 2, 6$. (b) 100 cu units. (c) $x = 4, z = 2$. $y = 3, z = 2$. $x = -1, z = 3$. **3.** (a) $6\sqrt{5} + 3\sqrt{6}$. (b) Area $\frac{49}{2}$. **5.** (a) $(0, 3, 0)$. (b) $(0, 2, 3)$. **9.** $x^2 + y^2 + z^2 = 10y$. **11.** (a) Center $(1, -2, -3)$, $r = 4$. (b) Point $(3, -4, -2)$. (c) Center $(4, 1, -2)$, $r = 5$. (d) No locus. (e) Center $(\frac{1}{2}, -\frac{2}{3}, \frac{7}{6})$ $r = \frac{\sqrt{69}}{6}$ (f) Point $(-\frac{2}{3}, -\frac{1}{3}, 0)$. **13.** (a) $\mathbf{i} + 4\mathbf{k}$. (b) $3\mathbf{i} - 6\mathbf{j} + 8\mathbf{k}$. **15.** $(-\frac{2}{7}, \frac{45}{7}, \frac{34}{7})$.

### § 18-2

**1.** (a) $\frac{4}{9}$. (b) $25/\sqrt{903}$. (c) $-5/3\sqrt{30}$. (d) $\frac{16}{21}$. **3.** (a) and (d) collinear; others not.

7. $\pm\frac{1}{\sqrt{6}}(\mathbf{i} - \mathbf{j} + 2\mathbf{k})$.
9. $-\frac{4}{21}(\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}) + \frac{1}{21}(46\mathbf{i} + 13\mathbf{j} - 68\mathbf{k})$.
11. $\cos^{-1}\frac{3}{7} \sim 64°37'$.

§ 18-3

1. (a) $2x - y + z + 8 = 0$. (b) $2x + y - 5z + 2 = 0$.
(c) $3x - 12y + 4z + 26 = 0$. (d) $x - 3y + 2z + 11 = 0$.
5. 9 units. 7. $y = 0, 2x = y, 2x + y - 4z = 0, 2x + y + 2z = 6$. 9. 45°. 11. 13.

§ 18-4

1. (a) $\frac{x-2}{1} = \frac{y+3}{\sqrt{2}} = z - 1$. (b) $x - z = 1$.
3. (a) $7y + 4z = 11$. (b) $7x - 5z = 2$. (c) $4x + 5y = 9$.
5. (a) $\frac{x-3}{2} = \frac{y+1}{-3} = \frac{z-2}{4}$. (b) $x = \frac{y-2}{2} = \frac{z+3}{3}$.
7. $31x - 354y - 185z + 1909 = 0$.
9. $6x - y - z = 8$. 11. $7x - 3y + z + 2 = 0$.
13. $(-\frac{1}{3}, -\frac{25}{3}, \frac{22}{3})$. 15. $(-\frac{1}{13}, \frac{10}{13}, -\frac{22}{13}), r = 3$.

§ 18-5

3. $(\mathbf{A} \times \mathbf{B})\cdot(\mathbf{C} \times \mathbf{D}) = 0$. 5. $13/\sqrt{69}$.

§ 18-6

1. (a) Elliptic cylinder parallel to $z$-axis. (b) Paraboloid of revolution about $y$-axis. (c) Circular cylinder, axis $x = 0, z = 1$. (d) Paraboloid of revolution about $z$-axis. (e) Two parallel planes. (f) Right circular cone, axis along $z$-axis, vertex at $z = 5$. (g) Parabolic cylinder parallel to $x$-axis. (h) Parabolic cylinder parallel to $z$-axis.
5. $4z = (y - 4)^2$.
7. (a) $x^2 + z^2 = 4y$, paraboloid.
(b) $y^4 = 4(x^2 + z^2)$.
(c) $4(x^2 + y^2) = 9(z - 2)^2$, cone.
(d) $9(x^2 + y^2) - 4z^2 = 36$, hyperboloid of one sheet.
(e) $16x^2 - 9(y^2 + z^2) = 144$, hyperboloid of two sheets.
(f) $4(z^2 + x^2) + (y - 4)^2 = 16$, ellipsoid.
9. (a) $2Z = X^2 - Y^2$. (b) Hyperbolic paraboloid. (c) Hyperbolas.

§ 18-7

**1.** (a) $1 : \sqrt{2} : 1$, $(4, \sqrt{2}, 0)$. (b) 28 units.
**3.** (a) $0 : 1 : -3$. (b) $6 : 3 : 1$. (c) $-6\mathbf{i} + 6\mathbf{j} + 2\mathbf{k}$.
**5.** (a) Radii 5. Planes $3z = \pm 4y$.
(b) $25(y - 3)z : -9xz : -16xy$. $0 : 3 : 4$.
(c) $x = \frac{5}{4}\sqrt{8z - z^2}$, $y = \frac{3}{4}z$.
**7.** (a) On cone $b^2(x^2 + y^2) = a^2z^2$.
(b) $\mathbf{V} = \omega\left(-\frac{a\pi}{2}\mathbf{i} + a\mathbf{j} + b\mathbf{k}\right)$ and $\omega(-a\mathbf{i} - a\pi\mathbf{j} + b\mathbf{k})$.

(c) $\cos\phi = \dfrac{\sqrt{a^2 + b^2}}{\sqrt{a^2\theta^2 + a^2 + b^2}} \to 0$ as $\theta \to \infty$.

**9.** (a) $4\pi a$. (b) $a\sqrt{2}\sinh 1$.

## *Review Problems, End of Chapter XVIII*

**1.** (1.20, 1.811). **3.** 0.6271. **5.** 1.475.

**11.** (b) $\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ a_1 & b_1 & c_1 & 0 \\ a_2 & b_2 & c_2 & 0 \end{vmatrix} = 0.$

**13.** $\dfrac{x}{a} + \dfrac{y}{b} + \dfrac{z}{c} = 1$.
**15.** $7x - 3y - 4z = 23$.
**17.** $7x + 2y = 50$, $x = 2z$.
**19.** Cone $9z^2 = 16(x^2 + y^2)$.
**21.** 0.

# CHAPTER XIX

§ 19-2

**1.** (a) $\dfrac{\partial f}{\partial x} = \dfrac{y}{2\sqrt{xy}} + y^3\cos xy^3 - 2xy\sin x^2y,$

$\dfrac{\partial f}{\partial y} = \dfrac{x}{2\sqrt{xy}} + 3xy^2\cos xy^3 - x^2\sin x^2y.$

(b) $\dfrac{\partial f}{\partial x} = y\,(\cos xy)\,e^{\sin xy} - \sin(x + y)e^{\cos(x+y)},$

$\dfrac{\partial f}{\partial y} = x\,(\cos xy)\,e^{\sin xy} - \sin(x + y)e^{\cos(x+y)}.$

(c) $\dfrac{\partial f}{\partial x} = 6x^2(x^3 - 2y) + \dfrac{2x - y}{2\sqrt{x^2 - xy}},$

$$\frac{\partial f}{\partial y} = -4(x^3 - 2y) - \frac{x}{2\sqrt{x^2 - xy}}.$$

(d) $\dfrac{\partial f}{\partial x} = 2x \tan^{-1}\dfrac{y}{x} - \dfrac{x^2y}{x^2 + y^2},$

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

(e) $\dfrac{\partial F}{\partial x} = \left(\dfrac{1}{y} - \dfrac{z}{x^2}\right)\sec^2\left(\dfrac{x}{y} + \dfrac{y}{z} + \dfrac{z}{x}\right),$

$$\frac{\partial F}{\partial y} = \left(\frac{1}{z} - \frac{x}{y^2}\right)\sec^2\left(\frac{x}{y} + \frac{y}{z} + \frac{z}{x}\right),$$

$$\frac{\partial F}{\partial z} = \left(\frac{1}{x} - \frac{y}{z^2}\right)\sec^2\left(\frac{x}{y} + \frac{y}{z} + \frac{z}{x}\right).$$

(*f*) $\dfrac{\partial G}{\partial r} = 2r \sin\theta \cos\phi - \dfrac{\sin\phi}{r^2},$

$$\frac{\partial G}{\partial \theta} = r^2 \cos\theta \cos\phi, \quad \frac{\partial G}{\partial \phi} = -r^2 \sin\theta \sin\phi + \frac{\cos\phi}{r}.$$

(g) $\dfrac{\partial F}{\partial a} = \dfrac{a - b\cos\theta}{(a^2 + b^2 - 2ab\cos\theta)^{1/2}}, \quad \dfrac{\partial F}{\partial b} = \dfrac{b - a\cos\theta}{(a^2 + b^2 - 2ab\cos\theta)^{1/2}},$

$$\frac{\partial F}{\partial \theta} = \frac{ab\sin\theta}{(a^2 + b^2 - 2ab\cos\theta)^{1/2}}.$$

**3.** $(x + y + z)^2$.

**7.** At $(1, 1, -2)$.

**9.** (a) $z + 2 = 2\left(\dfrac{x}{a} + \dfrac{y}{b}\right)$; $2b : 2a : -ab$.

(b) $45x - 100y + 24z + 650 = 0$; $45 : -100 : 24$.

(c) $5x - 4y - 4z + 17 = 0$; $5 : -4 : -4$.

(d) $x + 2y - 3z = 8$; $1 : 2 : -3$.

(e) $6x - 2y + 15z = 22$; $6 : -2 : 15$.

(f) $3x + 3y - 5z = 8$; $3 : 3 : -5$.

**11.** (a) 1000 man-hours ($x = 10$). (b) For $y = 4$, $z$ is a maximum when $x = 8$, so a change in $x$ from 8 decreases $z$. For $x = 8$, $\partial z/\partial y = 4$ at $y = 4$, so a small increase in $y$ increases $z$ by 4 tons per unit increase in $y$. (c) $-\frac{35}{22}$. If $x$ is increased a small amount, $y$ can be decreased by approximately $\frac{35}{22}$ times this amount.

## § 19-3

**1.** $c\,dx\,dy + a\,dy\,dz + b\,dz\,dx + dx\,dy\,dz$.

**5.** (a) $\dfrac{y(x^2\,dz + 2xz\,dx) - x^2z\,dy}{y^2}$.

(b) $\dfrac{dx}{x} - \dfrac{y\,dy + z\,dz}{y^2 + z^2}$.

(c) $\dfrac{z(x\,dy + y\,dx) - xy\,dz}{z^2 + x^2y^2}$.

(d) $\left(\dfrac{x}{y + z}\right)^{-1/2} \dfrac{(y + z)\,dx - x\,(dy + dz)}{2(y + z)^2}$.

(e) $\dfrac{3x^2\,dx - 3y^2\,dy - (xy\,dz + xz\,dy + yz\,dx)}{2\sqrt{x^3 - y^3 - xyz}}$.

(f) $-e^{xy} \sin xyz(xy\,dz + xz\,dy + yz\,dx) + e^{xy} \cos xyz(x\,dy + y\,dx)$.

**7.** 6.40 **9.** 1.8%. **11.** (a) $\frac{25}{24}$ ft. (b) $\frac{5}{8}$ ft.

## § 19-4

**7.** $\dfrac{\partial^2 w}{\partial x^2} + \dfrac{\partial^2 w}{\partial y^2} + \dfrac{\partial^2 w}{\partial z^2} = \dfrac{d^2 w}{dr^2} + \dfrac{2}{r}\dfrac{dw}{dr}$.

**9.** $F'(u)\left(\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2}\right) + F''(u)\left[\left(\dfrac{\partial f}{\partial x}\right)^2 + \left(\dfrac{\partial f}{\partial y}\right)^2\right]$.

## § 19-5

**1.** (a) $5x$, $-x + 7y$.

(b) $\dfrac{-1 + z(2y^2 - 3x^2)}{(1 + xyz)^2}$, $\dfrac{7 + z(3y^2 + 4x^2) + xy(y - x)}{(1 + xyz)^2}$.

(c) $\dfrac{s(1 + t^2)}{(x^2 + y^2 + z^2)^{1/2}}$, $\dfrac{s^2t}{(x^2 + y^2 + z^2)^{1/2}}$.

**3.** $\dfrac{\partial u}{\partial \phi} = 7$, $\dfrac{\partial^2 u}{\partial r\,\partial \phi} = \frac{15}{4} + \frac{17}{2}\sqrt{3}$.

**9.** (a) $\left(\dfrac{\partial A}{\partial P}\right)_x$. $\left(\dfrac{\partial A}{\partial y}\right)_x \left(\dfrac{\partial y}{\partial P}\right)_x$.

## § 19-6

**1.** $\sqrt{33}$. **3.** 864 sq ft. **5.** $a^3b^2/64$. **7.** (a) 18, at $x = 3$, $y = 2$, $z = \frac{3}{2}$. **9.** (a) $x = y = 6$, $z = 12$. (b) $x = 3$, $y = 6$, $z = 9$. **11.** (a) Saddle point at $(\frac{3}{2}, \frac{7}{2})$. (b) Rel. max. at $(2, -1)$. (c) Saddle point at $(0, 0)$, rel. min. at $(0, -1)$ and $(0, 2)$. (d) Rel. max. at $(-2, 1)$. (e) Rel. max. at $(\frac{5}{2}, \frac{4}{3})$, saddle points at $(3, 2)$, $(3, 0)$, $(\frac{3}{2}, 2)$. (f) Rel. max. at $(0, 0)$, saddle points at $(3, 3)$, $(-3, -3)$, $(1, -1)$, $(-1, 1)$.

**§ 19-7**

**1.** (a) $-\frac{3}{5}$. (b) $-4\sqrt{5}$. (c) $-\frac{33}{5}$. (d) 0. (e) $\frac{1}{9}$. (f) $\frac{20}{729}$. **3.** (a) $\dfrac{4 - 3\sqrt{3}}{50}$. (b) $\frac{5}{117}$. (c) $\dfrac{y}{(x^2 + y^2)^{3/2}}$. **5.** (a) $69\sqrt{14}/7$. (b) $-70\sqrt{2}$. (c) $6\sqrt{61}$.

**§ 19-8**

**1.** (a) $\dfrac{\partial u}{\partial x} = \dfrac{2vx}{u + v}, \dfrac{\partial v}{\partial y} = \dfrac{-1}{2(u + v)}$.

(b) $\dfrac{\partial u}{\partial x} = \dfrac{-2v}{4uv + 1}, \dfrac{\partial v}{\partial y} = \dfrac{2u}{4uv + 1}$.

(c) $\dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y} = e^{-u} \cos v$.

(d) $\dfrac{\partial u}{\partial x} = \dfrac{uy - 4vx}{2(u^2 + v^2)}, \dfrac{\partial v}{\partial y} = \dfrac{4uy - vx}{2(u^2 + v^2)}$.

**3.** (a) $-\dfrac{F_1(x_0, y_0)}{F_2(x_0, y_0)}$.

**5.** $$\frac{\partial f}{\partial y} = -\begin{vmatrix} F_2 & F_4 & F_5 \\ G_2 & G_4 & G_5 \\ H_2 & H_4 & H_5 \end{vmatrix} \div \begin{vmatrix} F_3 & F_4 & F_5 \\ G_3 & G_4 & G_5 \\ H_3 & H_4 & H_5 \end{vmatrix}.$$

**7.** $(a + b + c)^2$. **9.** $\dfrac{\partial^2 z}{\partial x\, \partial y} = \dfrac{z^2 - x}{(z^2 + x)^3}$. **11.** Farthest point is (0, −1, 7); nearest is (4, 7, 3).

## CHAPTER XX

**§ 20-2**

**1.** (a) −7. (b) 27. (c) $\frac{45}{2}$. **3.** (a) 45. (b) 10. (c) $3\pi/4$. (d) 16. (e) $486\sqrt{3}/35$. **5.** (a) $M = ca^4/6$. $\bar{x} = \bar{y} = 2a/5$. (b) $M = 2cb^4/3$. $\bar{x} = \bar{y} = 5b/8$. (c) $M = ca^3/6$. $\bar{x} = a/2, \bar{y} = a/4$. (d) $M = ca^2b/6$. $\bar{x} = a/2, \bar{y} = b/4$. (e) $M = ca^2b/4$. $\bar{x} = 8a/15, \bar{y} = 2b/3$. (f) $M = 4cab^2/15$. $\bar{x} = 5a/16, \bar{y} = 3b/7$.

**7.** (a) $(\frac{43}{70}, \frac{59}{70})$. (b) $(\frac{1}{2}, \frac{3}{5})$. (c) $(\frac{11}{10}, \frac{24}{75})$.

**9.** $\displaystyle\int_0^a dy \int_y^a \sqrt{a^2 - x^2}\, dx$.

**§ 20-3**

**1.** (a) $\frac{1}{2}Ma^2$. (b) $\frac{35}{24}Ma^2$. (c) $\dfrac{\pi}{2}Ma^2$. **3.** (a) $\pi a^3/6$. (b) $4\pi/3$. (c) $\dfrac{16a^3}{9}(3\pi - 4)$. (d) $32a^3/9$. **5.** (a) $\dfrac{5\pi - 8}{8}a^2$. (b) $\frac{2}{3}a^2$. **7.** $\dfrac{a^2}{4}(2\pi - 2 - 3\tan^{-1} 2)$. **9.** $I_x = \frac{1}{4}Ma^2$, $I_y = \frac{5}{4}Ma^2$. **11.** $5a^4/48$. **13.** The one through the center of mass.

§ 20-4

**1.** (a) $\dfrac{a^2 + 2b^2}{2b}\cdot$ (b) $a/\sqrt{2}$. **3.** $\frac{4}{3}$.

§ 20-5

**1.** (a) $\dfrac{\pi}{6b}\,[(b^2 + 4a^2)^{3/2} - b^3]$. (b) $a^2\left(\dfrac{\pi}{3} + 2\sqrt{3} - 4\right)\cdot$

(c) $\frac{1}{8}\log(2 + \sqrt{3})$. (d) $2\pi a^2\sqrt{2}$. (e) $2a^2(\pi - 2)$.

**5.** $\frac{2}{9}(20 - 3\pi)$. **7.** $\dfrac{\pi a^2}{2}\,\sqrt{2} - 1)$.

**9.** (b) $\frac{2}{3}Mr^2$. (c) $\frac{1}{2}Mr^2$. **11.** $\left(\dfrac{2a}{3}, \dfrac{4a}{3\pi}, \dfrac{4a}{3\pi}\right)\cdot$

§ 20-7

**1.** $M = \pi abc\sigma/6$, $\bar{x} = 3a/8$. **3.** $M/5$. **5.** (a) 25. (b) $\frac{15}{2}$. (c) 16.
**7.** (a) $M = \pi abc\sigma/2$, $z = c/3$.

(b) $\left(\dfrac{16a}{15\pi}, \dfrac{16b}{15\pi}, \dfrac{c}{3}\right)\cdot$

**9.** (a) $a^5/15$. (b) $a^3h^2/3$. (c) $\frac{6}{5}$.

§ 20-8

**3.** $\bar{z} = 3h/4$.
**5.** $2\pi\lambda\sigma(b + h - \sqrt{b^2 + h^2})$.

**7.** $2\pi\lambda\sigma h\left(1 - \dfrac{h}{\sqrt{a^2 + h^2}}\right)\cdot$

**9.** $\mathbf{F} = \dfrac{\pi\lambda\sigma ca}{\sqrt{1 + c^2}}\,\mathbf{i} + 4\lambda\sigma a\left(1 - \dfrac{1}{\sqrt{1 + c^2}}\right)\mathbf{k}$.

§ 20-9

**1.** (a) $\frac{2}{5}Ma^2$. (b) $(0, 0, 3a/8)$. (c) $\frac{45}{112}$.
**3.** $\bar{z} = \frac{3}{8}a(1 + \cos\alpha)$.

**5.** (b) $\dfrac{2\pi\lambda\sigma a}{3}(\sqrt{2} - 1)$. (c) $2\pi\lambda a\sigma(1 - \sqrt{2}/3)$.

**9.** (a) $ka^2\log(1 + \sqrt{2})$.

## CHAPTER XXI

§ 21-2

**1.** (a) $y = \dfrac{1}{2}\left(1 + \dfrac{C}{x^2}\right)\cdot$ (b) $y = C(x + 3) - 5$.

(c) $\tan y + \sec x = C$. (d) $3x - \sqrt{a^2 - y^2} = C$.
(e) $y^2 = C|\cos 2x|$.

**3.** (a) $y = 2\left(3\frac{x-4}{x+4}\right)^{1/8}$.

(b) $y = 2\left(3\frac{4-x}{4+x}\right)^{1/8}$.

(c) $y = -2\left(\frac{4-x}{4+x}\right)^{1/8}$.

**5.** (a) $\sin y = \log \sec x + C$.
(b) $y = \sin^{-1}(\frac{1}{2} + \log \sec x)$.

**7.** (a) $y = Cxe^{-x}$. (b) $y = xe^{2-x}$.

**9.** (a) $y^2 = \log \cos^2 x + C$.

(b) $y^2 = \frac{x^2}{2} + \log |x| + C$.

(c) $y^2 = \log |x| + C$.

(d) $y = \sqrt{1-x^2} - \log \frac{1+\sqrt{1-x^2}}{x} + C$.

## § 21-4

**1.** (a) $x^2 - y^2 = 2Cx$. (b) $(3x^2 + y^2)y = k^3$.

**3.** (a) $y = \frac{x}{\log |x| - C}$. (b) $x^2 + y^2 = C(y - x)$.

(c) $\log(4x^2 + y^2) + \tan^{-1}\frac{y}{2x} = C$.

**5.** (b) $r = Ce^{\theta}$. (c) $r = \frac{C}{1 + \cos \theta}$.

## § 21-5

**1.** (a) $y = 1 + Ce^{-x^2/2}$. (b) $y = x \sin x + C \sin x$.

(c) $y = \frac{x}{2}[(\log x)^2 + C]$. (d) $y = \frac{\sqrt{1+x^2}}{x} + \frac{C}{x}$.

(e) $y = x^2 + Cx^2e^{1/x}$.

**3.** (a) $i = E/R - (E/R - i_0)e^{-Rt/L}$.

(b) $i = \frac{E_0}{R^2 + \omega^2L^2}(R \sin \omega t - \omega L \cos \omega t) + \left(i_0 + \frac{E_0\omega L}{R^2 + \omega^2L^2}\right)e^{-Tt/L}$.

**5.** (a) $q = CE + (q_0 - CE)e^{-t/RC}$.

(b) $q = \dfrac{E_0C}{R^2C^2\omega^2 + 1}(RC\omega \sin \omega t + \cos \omega t) + \left(q_0 - \dfrac{E_0C}{R^2C^2\omega^2 + 1}\right)e^{-t/RC}$.

§ 21-6

**1.** (b) At $O$, heading in negative $x$-direction. (c) $2a/3v$ time units. **3.** (a) $2a/3$. (b) $3a^2x = 4y^3 - 6ay^2 + 2a^3$. **5.** $y = Cx^2$. **7.** The catenary.

**9.** (a) $y = \dfrac{5700}{19 + 56e^{-0.265t}}$.

(b) Over $222\frac{1}{2}$ million. (c) 300 million.

**11.** (a) $v = \sqrt{\dfrac{mg}{k}} \tanh \sqrt{\dfrac{gk}{m}}\, t$.

(b) $k = 0.08$. 27.88 sec.

**13.** (b) $T = \dfrac{\sigma g b}{1 + \mu^2}[2\mu e^{\mu\theta} - 2\mu \cos\theta + (1 - \mu^2)\sin\theta]$.

$$h = \frac{2\mu b}{1 + \mu^2}(1 + e^{\mu\pi}).$$

**15.** $\sqrt{5} - 1$ hours before noon.

§ 21-7

**1.** (a) $y = C_1x^2 + C_2$.
(b) $x^2 + y^2 = 1$ or $x^2 + (y - 2)^2 = 1$.
(c) $x^2 + y^2 = 2x$.
(d) $x^2 + (y + C_2)^2 = C_1^2$.

§ 21-8

**1.** $y = 3 + C_1x + C_2\left(1 + \dfrac{x}{2}\log\dfrac{1 - x}{1 + x}\right)$.

§ 21-9

**1.** (a) $y = 3e^{-2x} - e^{3x}$.
(b) $y = e^{3x}(3x - 1)$.
(c) $y = 2(\sin 3x + \cos 3x)$.
(d) $y = e^{-\pi/4}e^x(\sin x - \cos x)$
(e) $y = C_1 + C_2e^{-2x}$.
(f) $y = e^{-2x}(C_1 \cos 3x + C_2 \sin 3x)$.

**5.** $I = C_1 \exp\left[\dfrac{-Rt}{L - M}\right] + C_2 \exp\left[\dfrac{-Rt}{L + M}\right]$, where $\exp x = e^x$.

§ 21-10

1. (a) $\pi/5$. (b) 10. (c) $\pi/3$.

3. $x = \dfrac{E}{(a^2 - p^2)^2 + 4\lambda^2 p^2} [(a^2 - p^2) \sin pt - 2\lambda p \cos pt]$.

5. $x = \dfrac{E}{a^2 - \lambda^2 - p^2} e^{-\lambda t} \cos pt$.

7. (a) Transient $y = e^{-t/2RC} (A \cos \beta t + B \sin \beta t)$,

   where $\beta = \left(\dfrac{1}{LC} - \dfrac{1}{4R^2C^2}\right)^{1/2}$. Steady state $y = E/R$.

   (b) $y = \dfrac{E}{R} + e^{-t/2RC} (C_1 + C_2 t)$.

# INDEX