

FIRST COURSE IN CALCULUS

Hollis R. Cooley

WASHINGTON SQUARE COLLEGE
NEW YORK UNIVERSITY

JOHN WILEY & SONS, INC., NEW YORK
CHAPMAN & HALL, LTD., LONDON

49274

COPYRIGHT, 1950, 1951

BY

HOLLIS R. COOLEY

COPYRIGHT, 1954

BY

JOHN WILEY & SONS, INC.

All Rights Reserved

*This book or any part thereof must not
be reproduced in any form without
the written permission of the publisher.*

Library of Congress Catalog Card Number: 54-5804

PRINTED IN THE UNITED STATES OF AMERICA

Preface

The motive which led to the preparation of this book was a simple one. I simply wished to set forth the thinking by which the concepts and methods of calculus are developed, neither insulting the intelligence of the students by timidly evading vital issues nor burdening them with formal demonstrations that might deflect their thinking from the main current of ideas. The first aim is to instill insight. Although I consider ability to manipulate the symbolic language of mathematics to have little value if understanding is lacking, I believe that understanding is greatly enhanced by intelligent practice in manipulation. The two should go together, but some understanding should precede the manipulations.

The book is addressed to the students. It is not a mere exercise book; it is to be read. My own experience and that of other teachers indicates that it can be read by students and, for the most part, understood, previous to classroom explanations. Rigor is not a primary aim. Such rigor as will be found is the direct outgrowth of the attempt to present matters clearly and convincingly. Some readers have nevertheless expressed the opinion that the book is more rigorous than most textbooks in the undergraduate field.

The general plan of the book is as follows. The first chapter poses the main problems, and thus sets a goal, and at the same time offers a challenge by revealing the basic difficulties. In order to surmount these difficulties, the theory of limits is expounded in Chapter 2. These chapters and the seven that follow constitute an introduction to the main concepts and operations of calculus. In order that the student's mind may be as free as possible to assimilate the ideas that are characteristic of calculus, in these chapters all extraneous difficulties are minimized by employing only algebraic functions. The transcendental functions are introduced in Chapters 10 and 11, after the student has become familiar with the new ideas and operations of calculus. The remainder of the book is concerned with the further development of calculus and its applications.

I have tried to present each new idea singly, with adequate motivation and without extraneous details that might detract from the main

point. This purpose has largely dictated the order as well as the method of presentation. Concrete rate problems accompany the work in differentiation. Maxima and minima are treated carefully but without insistence upon needless tests. Inverse differentiation is taken immediately after differentiation, without any reference to integration. Then, after the definite integral has been introduced through the area problem, the student is prepared for the fundamental theorem, and in this sequence the meaning of the theorem is appreciated, as is its immense value. Differentiation and integration are extended to the trigonometric functions as a part of a study of the properties of these functions. The logarithmic function is defined by an integral, and its properties, as well as those of the exponential function, are derived from this integral. I am convinced that this is the most meaningful way to present the logarithmic function and perhaps the easiest. Three short chapters are devoted to improper integrals, approximate integration, and hyperbolic functions, respectively. The problem of approximating functions by polynomials leads to Taylor's theorem, which, in turn, is naturally followed by the representation of functions by series. Power series are discussed rather thoroughly, including their uniform convergence. Finally, differentiation and integration are extended to functions of several variables.

Presentations that lay some claim to novelty include the comparison of the orders of $\log x$, x^n , and a^x , without a dependence on l'Hospital's rule, and the approach to arc length through approximating tangents, a method that is slightly easier than the usual method and one that can be immediately generalized to the calculation of the areas of surfaces. I have tried to afford sufficient space to discuss applications meaningfully. In proving Taylor's theorem lucidity seemed more important than formal elegance.

In the first course in calculus most teachers and textbook writers shun the δ, ϵ -terminology. Fears regarding this usage are well justified when it is introduced in definitions and proofs without suitable preparation. Such terminology is natural and indispensable, however, in culminating statements involving limiting operations, wherein it is desired to express precisely what has already been expounded more vaguely. Once the point of this mode of expression is made clear to the student he sees that there is no other way of expressing the situation exactly, and he welcomes it. In this book some statements and proofs relating to basic matters are made in the δ, ϵ -form, although I have tried to avoid using it needlessly.

Throughout the early chapters, derivatives are denoted by D , or D_x , or by primes, e.g., y' or $f'(x)$, the Leibniz notation being reserved until differentials have been explained. Similarly, prior to the identification of inverse derivatives with integrals, inverse derivatives are denoted by D^{-1} . These notations are quite satisfactory and have not been dropped abruptly upon the introduction of the more common ones. In particular, $D^{-1}f(x)$ indicates the inverse derivative of $f(x)$ rather better than does $\int f(x) dx$.

Answers to many problems in the technique of differentiation and integration are included with the exercises. Answers to other odd-numbered exercises are printed at the end. Instead of the usual brief table of integrals, an even briefer section will be found at the end of the book, including a few of the most useful formulas with hints for integrating the most common types.

The author of a new book in calculus obviously owes much to the writers of earlier books in the field. I am deeply indebted, in particular, to Professor R. Courant, whose book has long been a source of inspiration and has guided me at many vital points in the writing of my own. In addition, I am especially grateful to the teachers at Washington Square College for their patience and counsel during the period when the book was used in a trial edition.

HOLLIS R. COOLEY

January, 1954

Downloaded from www.dbraulibrary.org.in

Contents

1 THE PROBLEMS OF CALCULUS, 1

Introduction, 1. Variables, Functional Relations, 2. The Nature of a Functional Relation, 2. Some Convenient Notations, 5. The Problems of Calculus, 9. The Tangent Problem, 9. The General Rate Problem, 11. Intuitive Solution of the Tangent Problem, 14. The Area Problem, 16. The Work Problem: Generalization of the Area Problem, 17. Summary, 19.

2 THE FUNDAMENTAL PROCESS OF CALCULUS, 20

The Solution of the Tangent Problem Re-Examined, 20. Sequences, 21. The Limit of a Sequence, 23. Remarks on Convergence, 25. Theorems on Limits, 25. Limits of Some Important Sequences, 27. Solution of the Tangent Problem, 30. Limit of a Function, 31. Examples. Theorems on Limits, 33. Continuous Functions, 36. Theorems on Continuous Functions, 37. Continuity of Algebraic Functions, 39. The Tangent Problem. Recapitulation, 41.

3 DERIVATIVES, 44.

Derivatives, 44. Derived Functions, 45. Notation, 46. Some Important Formulas, 47. Derivative of a Sum, 49. Derivatives of Products and Quotients, 52. Related Rates, 54. Compound Functions. The Chain Rule, 56. Inverse Forms of a Functional Relation, 59. Inverse Functions, 60. The Power Formula for Rational Exponents, 61. Functions Expressed Parametrically, 62. Functions Expressed Implicitly, 63. Summary, 65.

4 SECOND AND HIGHER ORDER DERIVATIVES, 68

Differentiation of a Derived Function, 68. Geometrical Meaning of the Second Derivative, 71. Rectilinear Motion, 74. Examples of Rectilinear Motion, 75. Higher Derivatives of Parametric Functions, 79. Higher Derivatives of Implicit Functions, 81.

5 MAXIMA AND MINIMA, 83

The Problem of a Maximum or a Minimum, 83. Maxima and Minima and the Tangent Problem, 84. Definitions and Theorems, 87. Applications, 91. Alternative Method, 99. Other Methods, 102.

6 THE INVERSE RATE PROBLEM, 104

Introduction, 104. Operations and Their Inverses, 104. Inverse Differentiation, 105. The General Power Formula, 110. Application to Motion in a Straight Line, 114. Curvilinear Motion, 116. Distance along a Curve, 120. Examples of Curvilinear Motion, 121. Finding Curves from the Slope Formula, 125. Differential Equations, 128.

7 THE AREA PROBLEM, 130

Restatement of the Problem, 130. The Area under the Graph of $y = x$, 131. The Symbol Σ , 132. A Fundamental Summation Formula, 133. Formulas for $\sum_{i=1}^n i^2$ and $\sum_{i=1}^n i^3$, 135. The Area under the Curve $y = x^2$, 137. Definition of Area under a Curve, 139. The Area under the Curve $y = 1/x^2$, 141. The Definite Integral, 144. Theorems on Integrals, 149. Need for a Better Method of Evaluating Integrals, 154.

8 EVALUATION OF INTEGRALS, 155

A Mean Value Theorem, 155. A Function Defined by an Integral, 156. The Fundamental Theorem of Calculus, 158. Application to the Evaluation of Integrals, 159. Primitive Functions and Integrals, 162.

9 SOME APPLICATIONS OF INTEGRALS, 164

Introduction, 164. Areas, 164. Volumes of Some Special Solids, 171. Volumes of Revolution, 173. Other Solids Having Slice Formulas, 178. Work, 182. Liquid Pressure on a Vertical Wall, 186.

10 TRIGONOMETRIC FUNCTIONS, 191

Introduction, 191. Radian Measure, 191. The Functional Relation $y = \sin x$, 193. The Continuity of $\sin x$, 196. Other Trigonometric Functions, 197. Limits of Some Discontinuous Trigonometric Functions, 198. Differentiation of the Sine Function, 202. Differentiation of Other Trigonometric Functions, 203. The Functional Relation $y = \sin^{-1} x$, 206. Other Inverse Trigonometric Functions, 208. Applications, 212. Inverse Differentiation, 217. The Differential Equation $y'' + a^2y = 0$, 221. Definite Integrals, 224.

11 LOGARITHMIC AND EXPONENTIAL FUNCTIONS, 227

Introduction, 227. The Function $\int_1^x t^{-1} dt$, 228. $L(x)$ and the Laws of Logarithms, 229. Further Properties of $L(x)$, 231. The Base of the Logarithmic Function $L(x) = \log x$, 232. Differentiation of Logarithmic Functions, 234. The Inverse of $L(x)$, 237. Other Exponential Functions, 240. Differentiation of Exponential Functions, 240. A Logarithmic Simplification, 242. The Inverse Derivative of $1/x$, 244. Inverse Derivatives of Exponential Functions, 247. The System of Functions $\int_1^x t^{n-1} dt$, 248. Comparison of $\log x$, x^n , and a^x , 251. The Differential Equation $D_x y = ay$, 253. Definite Integrals, 255.

12 DIFFERENTIALS. THE MEAN VALUE THEOREM, 258

The Problem of Approximating a Functional Value, 258. Differences and Differentials, 260. Approximations by Means of Differentials, 261. Differentiable Functions, 264. The Symbols dy/dx and d/dx , 265. Differential Formulas and Their Use, 266. The Mean Value Theorem for Derivatives, 270. Rolle's Theorem, 271. Proof of the Mean Value Theorem, 271. The Mean Value Theorem and Differential Approximations, 275. The Fundamental Theorem on Primitive Functions, 277.

13 TECHNIQUE OF INTEGRATION, 278

Introduction, 278. Change of Variable, 279. Integration Formulas, 282. A Quadratic Reduction, 285. Integration of Some Rational Fractions, 287. Examples of Integration by Partial Fractions, 289. Integration by Parts, 295. Integration of Trigonometric Functions, 299. Trigonometric Substitutions, 303. Other Substitutions, 307. Integral Tables. Reduction Formulas, 311. Nonintegrable Forms. Approximations, 316. Miscellaneous Exercises in Integration, 317.

14 GEOMETRICAL APPLICATIONS. RECTANGULAR COORDINATES, 319

The Tangent Problem and Curve Tracing, 319. Intersections of Two Curves, 320. Maximum, Minimum, and Inflection Points. Tests, 321. Irrational Algebraic Functions, 325. Curves Given by Parametric Equations, 327. Length of a Curve, 333. Arc Length as a Function, 337. Curvature, 338. Radius and Circle of Curvature, 341. The Area Problem Reviewed, 343. Volumes of Revolution. Shell Method, 346. Surfaces of Revolution, 350.

15 GEOMETRICAL APPLICATIONS. POLAR COORDINATES, 355

Introduction, 355. The Polar Coordinate System, 355. The Tangent Problem in Polar Coordinates, 357. Arc Length in Polar Coordinates, 365. Curvature in Polar Coordinates, 367. Area in Polar Coordinates, 369.

16 PHYSICAL APPLICATIONS, 373

Curvilinear Motion. Velocity in Polar Coordinates, 373. Curvilinear Motion. Acceleration, 379. Acceleration. Polar Coordinates, 385. A New Work Problem, 386. Total Force of Gravitational Attraction, 388. Moments, 393. Centroids, 400. Moment of Inertia, 404. Polar Moment of Inertia. Volumes of Revolution, 407.

17 IMPROPER INTEGRALS, 410

Extensions of the Definition of an Integral, 410. The Integral of a Function with a Finite Discontinuity, 411. The Integral of a Function with an Infinite Discontinuity, 415. An Integral over an Infinite Range, 419. Tests for Convergence of Improper Integrals, 422.

18 HYPERBOLIC FUNCTIONS, 426

Circular Functions, 426. Hyperbolic Functions, 428. Exponential Form, 430. Properties of the Hyperbolic Functions. Graphs, 431. Integral Formulas, 432. The Catenary, 433. Miscellaneous Exercises, 435.

19 APPROXIMATE INTEGRATION, 436

Introduction, 436. Trapezoidal Rule, 437. Simpson's Rule, 439. Errors in Approximate Integrals, 444.

20 THE APPROXIMATION OF FUNCTIONS BY POLYNOMIALS, 446

The Importance of Polynomial Approximations, 446. Examples of Polynomial Approximations, 446. Discussion of $R_n(x)$, 448. Approximation of Logarithms, 451. The General Problem, 452. Taylor's Theorem, 455. Numerical Approximations, 461. L'Hospital's Rule, 463. Summary, 469.

21 INFINITE SERIES OF CONSTANTS, 470

Introduction, 470. The Sum of an Infinite Series, 471. Examples of Infinite Series, 472. Two Theorems on Infinite Sequences, 475. Fundamental Convergence Tests for Series of Positive Terms, 477. Comparison with an Integral, 481. A Ratio Test and a Root Test, 484. Absolute Convergence, 487. Alternating Series, 488. The Representation of Numbers by Infinite Series, 491. Summary, 492.

22 REPRESENTATION OF FUNCTIONS BY POWER SERIES, 494

The Possibility of Expressing a Function by a Series, 494. The Interval of Convergence of a Power Series, 495. Examples of Power Series for Known Functions, 498. Taylor's Series, 500. Uniform Convergence of Power Series, 502. Theorems on Power Series, 503. Power Series and Polynomials, 506. Power Series Solutions of Differential Equations, 508. Limitations of Power Series, 511.

23 GRAPHICAL REPRESENTATION IN THREE DIMENSIONS, 512

The Place of Graphs in the Study of Functions, 512. Rectangular Coordinates in Three Dimensions, 512. Locus Problems in Three Dimensions, 514. Distance from a Coordinate Axis, 516. Distance between Two Points, 517. The Equation of a Sphere, 518. Surfaces of Revolution, 520. Some Special Surfaces, 521. Direction Cosines: Lines through the Origin, 524. Direction Cosines: Any Lines, 528. Planes and Their Equations, 529. Straight Lines, 532. Curves in Space, 537. Other Coordinate Systems, 541. Summary, 544.

24 FUNCTIONS OF SEVERAL VARIABLES—DIFFERENTIATION, 545

Functions of Two or More Variables, 545. Partial Derivatives, 547. Tangent Plane to a Surface, 553. Existence of a Tangent Plane, 554. Derivative in a Given Direction, 557. Maxima and Minima of Functions of Two Variables, 560. Differential of a Function of Several Variables, 563. Compound Functions. The Chain Rule, 566. Implicit Functions, 570. Curves and Surfaces Given in Implicit Form, 572. Space Curves, 574.

25 FUNCTIONS OF SEVERAL VARIABLES—INTEGRATION, 577

Volume under a Surface, 577. A Double Integral, 578. A Two-Dimensional Iterated Integral, 579. Integrals over Nonrectangular Domains, 583. Iterated Integrals in Polar Coordinates, 590. The Area of a Surface, 595. Physical Applications of Double Integrals, 598. An Integral over a Three-Dimensional Domain, 601. A Three-Dimensional Iterated Integral, 602. Cylindrical and Spherical Coordinates, 606. Applications of Triple Integrals, 611. Summary and Conclusion, 615.

INTEGRAL FORMULAS AND HINTS FOR THEIR USE, 616**ANSWERS TO ODD EXERCISES, 619****INDEX, 637**

The Problems of Calculus

1. Introduction. The student taking up calculus for the first time will find that the subject is based on concepts that are new to him. These concepts are not hard to grasp if they are developed gradually and studied thoughtfully. Undue haste to learn technique and gain practical results is a bar to understanding, and hence is self-defeating. Manipulation in the absence of understanding is mere busy work.

Calculus is a numerical subject; that is, the elements from which it is built are numbers, its concepts are numerical, its operations apply to numbers, and its results are numbers or conclusions about numbers. For this reason it can be applied to quantitative problems, a quantitative problem being simply one that can be formulated numerically, and a quantitative result being one that can be expressed by a number (strictly, a real number) or a numerical relation. Because of the power of its methods, calculus has great practical use in nearly all fields in which quantitative problems abound.

The main purpose of this chapter is to pose the problems that lead us naturally to formulate the concepts on which calculus is based and that guide us in devising methods for solving these problems as well as a vast number of others that grow out of them. Before doing this it is necessary that some technical terms and symbols be understood and agreed upon. First, we assume a knowledge of the real number system. In particular, we assume that the student is familiar with the idea of a one-to-one correspondence between the real numbers and the points of a line. Complex numbers will not enter into our work in any essential way. A more advanced book would attempt to deepen and broaden the student's knowledge of the real numbers at the outset, in order that the subtleties of a more rigorous development of calculus might be appreciated. The goal of the present book is understanding of concepts and methods and the ability to apply them to problems. Since insight, rather than rigor, is the principal aim, we shall often rely on intuitive conclusions. Many

proofs will be omitted. In special instances the student will be directed to sources where the missing proofs can be found.*

2. Variables, Functional Relations. If a quantity under consideration is not fixed, i.e., if it can change during a discussion, we represent it by a letter such as x and call this letter a *variable*. A *constant*, on the other hand, can have only one value. A constant may be a known number such as 2 or -10 , or it may be unknown and be represented by a letter such as a . We often use the last letters of the alphabet, such as x , y , and z , to denote variables, and the first letters of the alphabet, such as a , b , and c , to denote constants, although this custom is not strictly followed.

Many quantitative problems lead to a study of relations among two or more variables, called *functional relations*. A precise definition of a functional relation will be given in Art. 3. An important part of any science is the study of the functional relations among its variables. Not all of these relations are useful. The scientist must determine which ones are significant and deserve study. When he finds some functional relation to be important, it is of the utmost value to him to have an analysis of this relation, with as many results as can be derived from it. Mathematics furnishes such analyses, often in advance of their need. When some new relation is discovered that has not been previously treated, mathematicians set to work to study it as completely as possible. Often a single functional relation is found to have applications in many fields.

3. The Nature of a Functional Relation. Relations like the following are familiar to every student of elementary mathematics and are examples of functional relations.

1. $A = s^2$. This formula for the area of a square assigns a value to A for every value of s .

2. $C = 2\pi r$. This is the formula for the circumference of a circle of radius r .

3. $LW = 12$. This equation states the relation between the length and width of a rectangle of area 12. If W is given any positive value, a corresponding value of L is determined by the equation. Thus, if $W = 2$, $L = 6$; if $W = 3$, $L = 4$; etc.

4. $s = 16.1 \cdot t^2$. This equation gives the number of feet s that a body falls in t seconds if the acceleration of gravity is 32.2 ft per sec per sec.

* The principal reference for such material will be *Differential and Integral Calculus*, by R. Courant, 2 volumes, published by Interscience, New York. It is exceptionally readable and informative, and is of a somewhat more advanced level than the present book. The student is urged to become acquainted with it early. It will be designated in future references simply by *Courant*.

5. $y = +\sqrt{4-x^2}$. We may think of this equation as merely relating two mathematical variables x and y . Its graph is the semicircle in Fig. 1.

6.

y (year)	1900	1910	1920	1930	1940
P (population)	8913	10,442	14,675	19,271	20,563

This table of corresponding values of y and P expresses a relation between the number of the year and the population of a city. Unlike the equations in examples 1-5, it gives values of P for a limited number of values of y .

7. The graph of Fig. 2 represents a functional relation since it enables us to find a value of y for each value of x from 1 to 5.

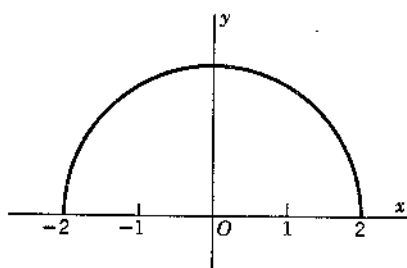


FIG. 1

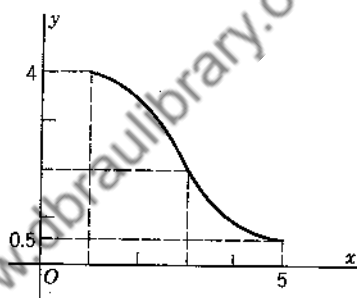


FIG. 2

A common property enjoyed by all these relations is the fact that for each value of one variable there corresponds a definite value of the other. In these examples the values the variables may assume are restricted, although this is not always true of functional relations. Thus, in Examples 1-4 we would permit only positive values of either variable; in 5 we would restrict x to values from -2 to $+2$, and then y would take on positive real values; in 6 there are just 5 values of each variable; in 7 the variable x can take on any value from 1 to 5, and the corresponding values of y lie between 0.5 and 4.

The definition of a functional relation between two variables can be stated as follows.

DEFINITION. The statement that there is a *functional relation* between x and y means that some rule exists whereby a value of one of these variables can be found corresponding to each value of the other. If the rule prescribes a value of y for each value of x , then y is said to be a *function of x* .

Note that, by this definition, a single value of one variable is made to correspond to each value of the other. We sometimes call attention to

this situation by saying that y is a *single-valued* function of x . If there were two values of y for each value of x , we would say that y is a *double-valued* function of x . In this book we shall be concerned entirely with single-valued functions, and any double-valued function will be treated as two single-valued functions. For example, the equation $y^2 = 4 - x^2$, or $y = \pm \sqrt{4 - x^2}$, assigns two values of y to each x between -2 and $+2$. We can, with greater clarity, however, treat the two relations $y = +\sqrt{4 - x^2}$ and $y = -\sqrt{4 - x^2}$ as separate single-valued functions. Usually, both quantities involved in a functional relation are variables. It is possible, however, that for all values of x , y is equal to the same quantity. In this case y is constant. When x and y are both variables, as is usual, it is often convenient to distinguish between the variables by calling the variable to which values are assigned at will the *independent* variable, and that whose values are determined by the rule the *dependent* variable. Thus, if, for each value of x , the rule assigns a value of y , so that y is a function of x , then y is the dependent variable and x is the independent variable. It is clear that often it is a matter of arbitrary choice which variable shall be considered independent, since a rule that gives a value of y for each value of x can often be restated so as to give a value of x for each value of y .

In order that y shall be defined clearly as a function of x , two things are necessary. First, we must know what values of the independent variable x are permitted. The whole set of such values is called the *range* of x . The range may be the set of all real numbers; it may consist of all real numbers between two fixed limits, a and b ; or it may be further restricted, perhaps to a definite set of specified numbers.

Second, we must have a rule for determining a value of the dependent variable y for any given value of x . This rule can be stated in any way that provides unmistakable directions for associating a value of y with a given value of x . Thus, we might let x be the number of any page in the Manhattan telephone directory. The rule for finding corresponding values of y might be: turn to the page numbered x and select the last digit of the tenth telephone number in the first column. This is a perfectly clear rule that anyone could follow. In a similar way, a table of logarithms defines a functional relation between numbers and their logarithms. If x is a number and y is its logarithm, the rule for finding y is the rule for looking up a logarithm in the table. Likewise, any tabulation of corresponding numbers defines a functional relation. Thus, in Example 6 above, we think of P as a function of y .

The two most convenient forms that the statements of functional relations can take are equations and graphs. The use of equations for this purpose is illustrated by Examples 1-5 above. E.g., in Example 1,

A is a function of s . Graphical statement of functional relations is illustrated in Examples 5 and 7, in each case y being a function of x . Although by far the most common graphical representation of functional relations employs a rectangular coordinate system, other types of graphs have special advantages, particularly graphical representation on a polar coordinate system. Polar representation will be discussed in some detail in a later chapter.

Since the invention of coordinate geometry by Descartes, in 1637, graphs have played an important role in the study of relations among variables. Indeed, functional relations and their graphical representations are identified to such an extent that we often use the word "point" as an abbreviation for "coordinates of a point." It is, of course, the numerical coordinates that are related. We shall see that the problems of calculus that we are about to consider are closely connected with geometric problems, and that the visual appeal of graphical representations will help us to find the desired solutions.

It is important to notice that a functional relation implies nothing more than a correspondence between the values of two or more variables. The telephone-book function defined above is this kind of correspondence. It is meaningless to raise a question as to a causal relation among the variables. Physical considerations will often determine what relations will be fruitful for study. The mathematical relation, however, is examined without bias arising from any sources, as a numerical relation and nothing more.

A functional relation among more than two variables is defined in a way similar to that for two. We define it for three variables.

DEFINITION. The statement that there is a functional relation among x , y , and z means that some rule exists whereby a value of one of these quantities can be found corresponding to each pair of values of the other two. If the rule prescribes a value of z for each determination of x and y , then z is said to be a function of x and y . As with a relation between two variables, a clear definition of z as a function of x and y requires that the range of values that may be assigned to x and y be specified.

4. Some Convenient Notations. It is conventional to express the fact that y is a function of x by a symbolic equation like $y = f(x)$. This is read " y is a function of x " or " y equals f of x ." Although the word "function" suggests the letter f , there is no reason why this letter must be used. When a letter has been adopted to denote a given function, however, that letter must denote the same functional relation throughout the discussion. If another function is to be considered at the same time, it must be denoted by a different letter. Thus, if the length of a side of a square is x , the area and perimeter are $A = x^2$ and $P = 4x$, respectively.

Here we have two functions of x . In the functional notation just given we might write $A = f(x)$ and $P = g(x)$. Greek letters, like ϕ (phi) and ψ (psi), are often used with this notation, i.e., $\phi(x)$ and $\psi(x)$.

The above notation is often used to indicate that a functional relation exists, even though we may not know its precise nature. It is also used to abbreviate a known relation when it is unnecessary or inconvenient to state the relation explicitly. Thus, we can let $y = f(x)$ stand for the functional relation expressed by a complicated equation, one expressed by a graph, or one expressed by a table such as the table of logarithms. To say that a value of $f(x)$ is determined for each value of x is, of course, equivalent to saying that a value of y is determined for each value of x .

The notation $y = f(x)$ is often applied to fixed values of the independent variable. For example, if

$$y = f(x),$$

in which

$$f(x) = x^2 - 5x + 3,$$

then, corresponding to $x = 1$, we have

$$y = f(1) = 1^2 - 5 \cdot 1 + 3 = -1;$$

corresponding to $x = 2$, we have

$$y = f(2) = 2^2 - 5 \cdot 2 + 3 = -3;$$

corresponding to $x = -1$, we have

$$y = f(-1) = (-1)^2 - 5 \cdot (-1) + 3 = 9; \text{ etc.}$$

The symbol $f(a)$ denotes the value obtained by replacing x by a in $f(x)$, i.e., the value assumed by $f(x)$ when $x = a$. One may think of $f(\quad)$ as standing for a rule for determining a functional value corresponding to any number one may insert in the parentheses. In the above example, $f(\quad) = (\quad)^2 - 5(\quad) + 3$. Thus

$$f(0) = (0)^2 - 5 \cdot (0) + 3 = 3;$$

$$f(a) = (a)^2 - 5 \cdot (a) + 3;$$

$$f\left(\frac{2}{1-x}\right) = \left(\frac{2}{1-x}\right)^2 - 5 \cdot \left(\frac{2}{1-x}\right) + 3.$$

We see that $f(x)$ can replace y in an equation like $y = x^2 - 5x + 3$, obtaining $f(x) = x^2 - 5x + 3$, without changing the meaning of the equation.

Except for the last three chapters of the book, we shall be concerned, for the most part, with functional relations between only two variables. Occasionally, more variables will be involved. Thus, the area A of a

rectangle is related to the length L and the width W by the equation $A = LW$. Here we think of A as the dependent variable and both L and W as independent variables. We say that A is a function of L and W . The general notation indicating that u is a function of two independent variables x and y is $u = f(x, y)$. Remarks made above about the functional symbol $f(x)$ apply equally to $f(x, y)$ and to symbols for functions of more variables. Thus, for example, if $f(x, y) = x + y - \sqrt{xy}$, then $f(1, 4) = 1 + 4 - \sqrt{1 \cdot 4} = 3$. That is, if $x = 1$ and $y = 4$, then $f(x, y) = 3$.

Similarly, we might have

$$F(x, y, z) = x^2 + y^2 + z^2 - 3xyz,$$

whence,

$$F(2, 5, -1) = 2^2 + 5^2 + (-1)^2 - 3 \cdot 2 \cdot 5 \cdot (-1) = 60.$$

In order to define a specific function, $f(x)$, it is not sufficient merely to state $f(x)$ explicitly. In addition, as stated in Art. 3, we must give the range of values permitted the variable x . In practice, the range is often taken for granted, particularly when it includes all real numbers; but if there is any chance of misunderstanding, the range must be specified. If the range is an interval, i.e., if it includes all real numbers between two fixed numbers a and b , with $a < b$, this fact can be expressed by the inequality $a < x < b$, which is read " a less than x less than b ." This interval is *open* since the end-points a and b are not included; i.e., x cannot equal a or b . If we want to include a and b , and thus make the interval *closed*, we write $a \leq x \leq b$ ($x \leq b$ means that x is less than or equal to b).

Many discussions in calculus are simplified by the use of *absolute values* of numerical expressions. The absolute value of an expression representing a real number is its arithmetic, or unsigned, value as opposed to the algebraic, or signed, value. The absolute value of a number is denoted by placing vertical bars around it. Thus, the absolute value of -5 is denoted by $|-5|$ and is equal to the arithmetic number 5. Similarly, $|-1| = |1| = 1$; $|1 - 1.2| = |1.2 - 1| = .2$; and $|x - a| = |a - x|$. The following theorems governing inequalities involving absolute values of real numbers are easily verified:

$$\begin{aligned} |a + b| &\leq |a| + |b|, & \text{e.g.,} & & |2 + 3| &= |2| + |3| \\ & & \text{and} & & |2 + (-3)| &< |2| + |-3|; \end{aligned}$$

$$\begin{aligned} |a - b| &\geq |a| - |b|, & \text{e.g.,} & & |2 - 3| &> |2| - |3| = -1 \\ & & \text{and} & & |3 - 2| &= |3| - |2| = 1; \end{aligned}$$

$$|ab| = |a| \cdot |b|, \quad \text{e.g.,} \quad |2 \cdot (-3)| = |2| \cdot |-3| = 6.$$

The student should verify the fact that each of the following statements completely defines a function. Here, as elsewhere in this book, each letter stands for a real number unless otherwise specified.

$$1. f(x) = x^2 + 2, x \geq 0.$$

$$2. f(x) = \sqrt{1-x}, |x| \leq 1.$$

$$3. g(x) = 1/x, x \neq 0.$$

$$4. F(x) = |x|, x = \text{any real number}.$$

5. $\phi(x) = x$, if $x < 0$; and $\phi(x) = x^2$, if $x \geq 0$. Note that the two formulas for determining $\phi(x)$ apply to different parts of the range of x , which is all real numbers. These two formulas constitute one unambiguous rule for finding $\phi(x)$; i.e., a single definite value of $\phi(x)$ is found for each x .

$$6. f(x) = 1 \text{ for all real values of } x.$$

7. $f(x) = 1$, if x is an integer; and $f(x) = \frac{1}{2}$, if x is not an integer, the range being all real numbers.

$$8. F(x) = \frac{x^2 - 4}{x - 2}, x \neq 2 \\ = 5, x = 2.$$

$$9. \phi(x) = \log x, x > 0.$$

$$10. f(x) = \frac{10^x - 1}{x}, x \neq 0.$$

EXERCISES

1. What is the functional relation between the circumference C and the radius R of a circle? Find C when $R = 3$. How does C change when R changes 1 unit? If $R = f(C)$, write $f(C)$ explicitly. Find R when $C = 25$.

2. What is the functional relation between the area A and the radius R of a circle? If $R = f(A)$, write $f(A)$ explicitly. Find R when $A = 36$.

3. Express the side x of a square as a function of the area A . Express A as a function of x . Can you tell how A changes relative to a unit change in x ? How does this rate of change compare when $x = 2$ and when $x = 10$?

4. If r is the radius of a sphere, V is its volume, and S is the area of its surface, we have $V = (\frac{4}{3})\pi r^3$, and $S = 4\pi r^2$. Hence, $V = f(S)$. Write $f(S)$ explicitly.

5. If an airplane flies at the constant speed of 300 mph, write s as a function of t , if s denotes miles and t denotes hours.

6. Given $f(x) = 1 + 7x - 3x^2$, find $f(0)$, $f(2)$, $f(10)$, $f(-3)$, $f(\frac{1}{2})$, $f(a)$, $f(2a)$, $f(-a)$, $f(-x)$, $f(a+b)$, $f(a) + f(b)$, $f(y^2)$, $f(a/b)$, $f(a)/f(b)$.

7. Draw the graphs of the following functions:

$$(a) y = x^2, x \leq 0.$$

$$(c) y = 1/x, 1 \leq x \leq 10.$$

$$(b) y = x^2, -2 \leq x \leq 2.$$

$$(d) y = |x|, -2 \leq x \leq 2.$$

$$(e) y = |1 - x|, -2 \leq x \leq 3.$$

$$(f) y = -1, x < 0 \\ = x^2 - 1, x \geq 0.$$

$$(g) y = -\sqrt{4 - x^2}, |x| < 2 \\ 0, |x| > 2.$$

$$(h) y = 0, x \neq \text{an integer} \\ = 1, x = \text{an integer}.$$

8. In each of the following exercises use a table to find the indicated functional values, and draw the graph of $y = f(x)$ in the given range.

$$(a) f(x) = x \log_{10} x, 0 < x \leq 2; f(2), f(1), f(.5), f(.1), f(.01), f(.001).$$

$$(b) f(x) = \frac{\log_{10} x}{x - 1}, 0 < x < 1 \text{ and } 1 < x \leq 2; f(2), f(1.5), f(1.1), f(.9), f(.5), f(.1).$$

$$(c) f(x) = \frac{10^x - 1}{x}, -1 \leq x < 0 \text{ and } 0 < x \leq 1; f(-1), f(-.5), f(-.1), f(.1), f(.5), f(1).$$

$$(d) f(x) = \frac{\sin x}{x}, 0 < x \leq \pi/2; f(\pi/2), f(\pi/3), f(\pi/6), f(\pi/90).$$

9. The symbolic equations (1) $F(x + y) = F(x) \cdot F(y)$, (2) $F(x) + F(y) = F(xy)$, and (3) $F(x + y) = \frac{F(x) + F(y)}{1 + F(x) \cdot F(y)}$ are examples of *functional equations*, i.e., equations in which the unknown is the form of a function. Thus, (1) holds if $F(x) = a^x$. Find a function satisfying (2); and another satisfying (3).

10. If $f(x) = x^2$, find $f(x + h)$ and $f(x + h) - f(x)$.

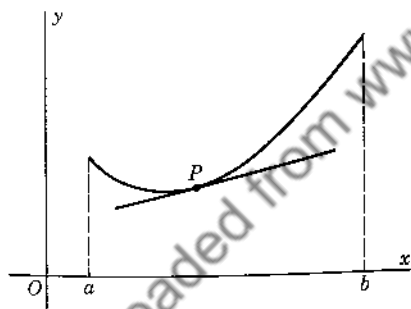


FIG. 3

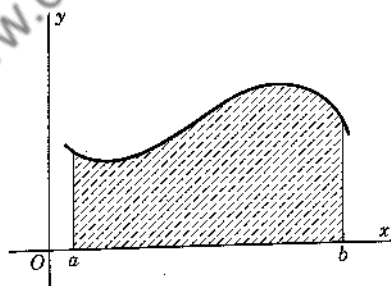


FIG. 4

5. The Problems of Calculus. The processes of calculus yield the solutions of a great variety of problems, and the basic ideas of calculus throw light on many of the concepts of science. All these problems and concepts, however, are closely related to one or the other of the two basic problems of calculus. These two problems arise again and again in various forms. In their geometrical forms they are the following.

1. The problem of determining the slope of the tangent to a given curve at a specified point, from the knowledge of the coordinates of the points on the curve in the vicinity of this point. In Fig. 3, for example, if the equation of the curve is $y = f(x)$, where $f(x)$ is a known expression, we wish to find the slope of the tangent at a point P .

2. The problem of evaluating the area bounded by the x -axis, the two ordinates $x = a$ and $x = b$, and a curve $y = f(x)$, where $f(x)$ is a known expression that has a positive value for every value of x in the interval $a \leq x \leq b$ (Fig. 4).

We shall examine these problems in more detail in the remainder of this chapter, with the purpose of showing something of what their solutions demand and where the difficulties lie. We shall also obtain the solution of the tangent problem in simple examples, proceeding intuitively.

6. The Tangent Problem. Suppose we are given the equation of a straight line in rectangular coordinates, e.g., $x - 3y + 5 = 0$. The

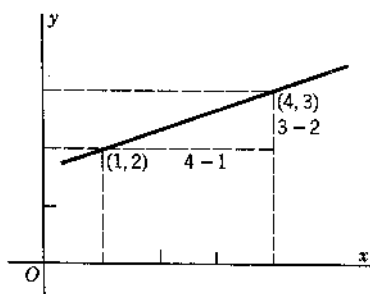


FIG. 5

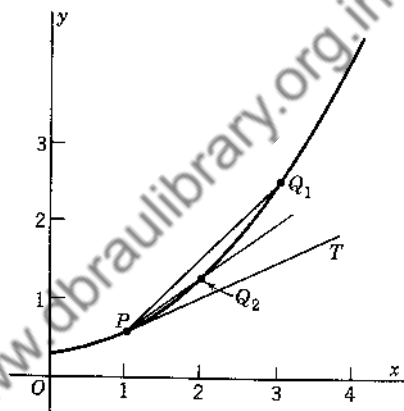


FIG. 6

direction of this line is expressed by means of the slope. This is constant and easily found from any two points of the line. Thus, in Fig. 5, by use of the points $(1, 2)$ and $(4, 3)$, the slope is found to be $\frac{3 - 2}{4 - 1} = \frac{1}{3}$.

It is not so simple as this to express the direction of a curve (Fig. 6). One naturally thinks that the direction of the curve as it passes through P is the same as that of the tangent at P . But what is this direction? Without further knowledge we cannot even draw a tangent accurately.

Although we cannot immediately find the direction of the tangent to a curve at a given point, as we found the direction of a straight line, we can approximate it by the same means. Thus, from the coordinates of P and those of another point Q on the curve, we can easily find the slope of the chord PQ . Of course, this is usually quite different from the slope of the tangent at P . If several chords are drawn through P , however, as shown in Fig. 6, with end-points Q_1, Q_2 , etc., closer and closer to P , it appears that the direction of the chord approximates that of the tangent more and more closely. Let us try, in this way, to approximate the slope of the

tangent to the curve whose equation is $y = x^2$ at the point $P(1,1)$. The coordinates of Q_1 are $x = 2, y = 4$. Hence, the slope of PQ_1 is $(4 - 1)/(2 - 1) = 3$. The coordinates of Q_2 are $x = \frac{3}{2}, y = \frac{9}{4}$. Hence, the slope of PQ_2 is $(\frac{9}{4} - 1)/(\frac{3}{2} - 1) = \frac{5}{2}$. If we obtain further approximations by using chords through P and Q_3 , where $x = \frac{4}{3}$; P and Q_4 , where $x = \frac{5}{4}$; etc., we get the approximating slopes $\frac{7}{3}, \frac{9}{4}$, etc. The figure indicates that the values thus obtained would approximate the slope of the tangent more and more closely as x decreases toward 1. We shall see later that the slope of the tangent is actually 2.

Perhaps it is time to point out that as yet we not only have no means of finding the slope of the tangent exactly, but it has not been made clear what a tangent really is. What we have done is to find approximations to the slope of a line that we accept intuitively as the tangent. The steps, in order, by which we shall pursue the problem further are the following: we shall devise a method of finding the precise value that we are now approximating; we shall then define the slope of the tangent to be the result that this process yields; the tangent at P will be defined as the line through P having this slope; and, finally, the slope of the curve at P will be defined as the slope of the tangent through P , as appeared obvious at the start.

EXERCISES

1. Find two approximations to the slope of the tangent to the curve $y = x^2 - x + 8$ at $P(1, 8)$, by the use of $Q_1(2, 10)$ and $Q_2(\frac{3}{2}, \frac{25}{4})$. Draw the curve and the chords PQ_1 and PQ_2 , and estimate the slope of the tangent.
2. Find two approximations to the slope of the tangent to the curve $y = x^2/4 - x + 2$ at $P(2, 1)$, by the use of $Q_1(4, 2)$ and $Q_2(3, \frac{5}{4})$. Draw the curve and the chords PQ_1 and PQ_2 , and estimate the slope of the tangent.

7. The General Rate Problem. Among the problems that are closely related to that of the slope of a curve is the problem of finding the speed of a body in motion from the functional relation between distance and time. The latter quantities are two elementary physical variables that are measured by familiar means. The distance, s units, that a body moves in t units of time is often expressed by an equation of the form $s = f(t)$. The simplest case is that of uniform motion, in which a body moves through equal distances in equal times, i.e., $s = kt$. Here we define the speed as the ratio of the number of units of distance moved to the number of units of time required. That is, if the body moves s units of distance in t units of time, the constant ratio, $s/t = k$, is called the speed of the body. The notion of speed is a derived one, arising from a relation between distance and time.

This definition of speed is satisfactory as long as we are dealing with

uniform motion. But suppose the motion is nonuniform, i.e., it is being accelerated or retarded. The definition no longer seems appropriate. For example, a falling body may pass a certain point P ; in the next second it is observed to fall 48 ft; if the speed is taken to be the ratio of the distance to the time, we would say that the speed is $48/1$, or 48, ft per sec. But if we observe it for 2 sec instead of 1, we find that it has fallen 128 ft. The ratio of distance to time is here $128/2$, or 64. Which is right, if either is? And what does the answer mean? Is it the speed at the beginning of the time of observation, at the end, in the middle, or during the entire time? The most we can say is that 48 ft per sec is the speed the body would have to maintain during the first second of observation, if the motion were uniform for this second, in order to fall the distance it is observed to fall. Similarly, 64 ft per sec is the speed the body would have to maintain during the 2 sec of observation, if the motion were uniform for this time, in order to fall the distance it really does fall in 2 sec. These numbers are, therefore, fictitious speeds, or substitute speeds—uniform speeds that would move the body the observed distances if the fictitious motions could be substituted for the real ones. They are often called *average* speeds; i.e., 48 ft per sec is the average speed for the first second, and 64 ft per sec is the average speed for the 2 sec.

If we find a substitute speed for a very small interval of time, it will not differ greatly from the true speed, because the true speed will not change much in this short time. Obviously, by considering shorter and shorter intervals, we can obtain substitute speeds closer and closer to the true speed. By this time, surely, one must feel compelled to ask what is meant by "true speed." Although this term has not been defined, we feel intuitively the existence of a speed more intimately bound up with the motion than any of the substitute speeds. We feel that there is some kind of speed that is connected with each point of the motion, i.e., with each instant during the motion. Such a speed we might call an *instantaneous speed*.

The similarity of this problem to that of the slope can hardly be missed. In the two problems, the approximations—slope of the chord in one and substitute speed in the other—are found in precisely the same way. The two problems can be made identical simply by plotting the graph of the relation between distance and time. Letting values of s be represented parallel to the y -axis, and those of t parallel to the x -axis, the graph of a uniform motion would be represented by a straight line, as in Fig. 7, in which the slope of the line represents the speed. If the motion is accelerated, the values of s increase more and more rapidly as t increases; i.e., the graph is a curve rising more and more rapidly as it proceeds toward the right, as in Fig. 8. The substitute speed for an interval of time from

t_0 to t_1 is obtained by dividing the number of units of distance moved in this time, i.e., $s_1 - s_0$, by the duration of the interval, $t_1 - t_0$, and is represented by the slope of the chord PQ_1 . Reasoning intuitively, we would expect the speed at the instant $t = t_0$ to be the same as the slope of the tangent at P . The two problems are therefore identical in character.

The tangent problem can be identified with a much larger class of problems. Suppose two variables u and v are related by an equation $v = ku$, where k is a constant. Clearly, any change in u produces a proportional change in v . This kind of relation is similar to that between

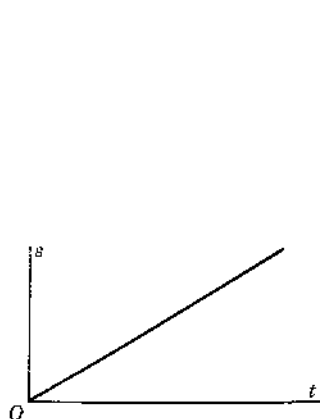


FIG. 7

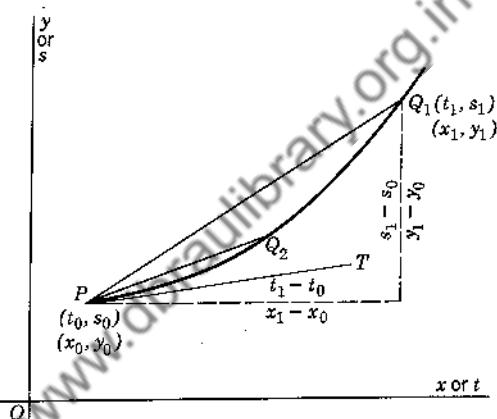


FIG. 8

s and t in uniform motion. The graph of the relation is a straight line, and the slope of this line is equal to the constant k , which is the ratio of the change in v to that in u . We call this ratio the "rate of change of v per unit change of u ," or the "rate of change of v with respect to u ." Thus, since the circumference of a circle is $C = 2\pi r$, we find that C changes 2π units per unit change in r .

Now, suppose u and v are related in a more complicated way, so that the graph of $v = f(u)$ is no longer a straight line. How shall we find the rate of change of v per unit change of u ? We can find a substitute rate, as with speed, but we are really interested in finding a precise rate associated with a definite pair of values of u and v , for that is the sort of thing we have in mind when we speak of the rate of change of v per unit change of u . The problem will be solved by finding the slope of the tangent of the curve whose equation is $v = f(u)$, at the point on the curve corresponding to the definite values of u and v which we have in mind. The tangent problem is thus a geometrical interpretation of the problem of finding the rate of change of one variable with respect to another; in fact,

it is the same problem expressed in geometrical language. When we discover a method of solving the tangent problem, we shall thus have, at the same time, a method of solving a great variety of rate problems.

Some examples of rates that occur in practical matters are the following.

The rate of rise of a hill, called the grade, or slope of the hill.

The rate of change in the amount of salt that can be dissolved in water per unit change in temperature.

The rate of change in the pressure in the ocean per unit increase in depth.

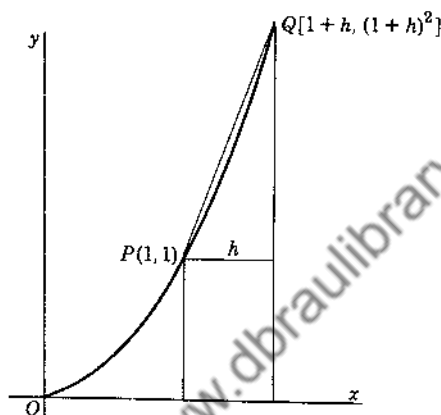


FIG. 9

The rate of change in the cost of operating an automobile with respect to a change of the speed at which it is driven.

The rate of increase in the resistance of the air to the motion of an airplane per unit increase in the speed.

The rate of change in the profits of a manufacturing concern per unit change in output.

The rate of increase in a bacteria culture per unit of time.

The rate of change in the rate of growth of a bacteria culture per unit change in temperature.

8. Intuitive Solution of the Tangent Problem. In simple cases, we can obtain a solution of the tangent problem that is intuitively reasonable by generalizing the approximation process outlined in Art. 6. Let us try to find the precise slope of the tangent to the parabola, $y = x^2$, at the point $P(1, 1)$ (Fig. 9). We draw a chord through P and a second point Q at which the abscissa differs from 1 by an unspecified amount h . The coordinates of Q are then $(1 + h, (1 + h)^2)$ or $(1 + h, 1 + 2h + h^2)$, and the slope of PQ is

$$\frac{1 + 2h + h^2 - 1}{h} = \frac{2h + h^2}{h} = 2 + h.$$

Hence, by considering chords through P and a second point Q , with Q coming closer and closer to P and h becoming smaller and smaller, we see that the slopes of these chords are closer and closer to 2. Thus, the slope that is approximated better and better by the slopes of these chords appears to be 2, and we readily accept 2 as the slope of the tangent to the curve at P . Although "slope of the tangent to a curve" has not yet been defined, at a later point we shall define it in a way that is consistent with this intuitive result.

We can generalize this result by taking P to be any point of the parabola, in place of $(1, 1)$. Let the abscissa of P be x_1 . Its ordinate must then be $y_1 = x_1^2$. Now, take Q at $x = x_1 + h$; then the ordinate of Q is $(x_1 + h)^2$. As before, the slope of PQ is

$$\frac{(x_1 + h)^2 - x_1^2}{h} = \frac{2hx_1 + h^2}{h} = 2x_1 + h.$$

Hence, as Q takes positions on the curve nearer and nearer to P , i.e., with h smaller and smaller, the slope of PQ tends more and more to $2x_1$. We conclude that the slope of this parabola at any point is twice the abscissa of the point. Denoting the slope by m , we can write: the formula for the slope of the parabola $y = x^2$, at any point, is $m = 2x$.

As a second example, let us try to find the slope of the hyperbola $y = 12/x$ at $P(3, 4)$. As before, we consider a second point Q , with abscissa differing from 3 by an unspecified amount h : $Q(3 + h, 12/(3 + h))$. The slope of PQ is then

$$\frac{12/(3 + h) - 4}{h} = \frac{-4h}{h(3 + h)} = -\frac{4}{3 + h}.$$

As Q is taken closer and closer to P , on the curve, i.e., h has a smaller and smaller value, this fraction tends more and more to $-4/3$. Hence, we conclude that the slope of the hyperbola at $(3, 4)$ is $-4/3$. The student can carry out this process for a general point of the hyperbola, $P(x_1, 12/x_1)$, with Q having the coordinates $(x_1 + h, 12/(x_1 + h))$ (see Exercise 3). As a result, he will obtain $m = -12/x_1^2$, and hence conclude that the formula for the slope of the hyperbola $y = 12/x$, at any point, is $m = -12/x^2$.

The intuitive method of this article points in the direction of the complete solution of the tangent problem. The step by which we got the final result, however, had the nature of jumping to a conclusion. Since this step is basic in all the work of calculus, it must be examined carefully. This we shall do in Chapter 2.

EXERCISES

1. From the formula for the slope of the parabola $y = x^2$, find the slope at $(0, 0)$, $(1/2, 1/4)$, $(2, 4)$, $(-1, 1)$, and $(-2, 4)$.
2. Verify the slope of $y = x^2$ at $(-2, 4)$ obtained by formula, by finding the slope directly, as was done for the point $(1, 1)$.
3. Obtain the general formula for the slope of the hyperbola $y = 12/x$.
4. From the formula for the slope of the hyperbola $y = 12/x$, find the slope at $(2, 6)$, $(4, 3)$, $(6, 2)$, and $(-3, -4)$.
5. Verify the slope of $y = 12/x$ at $(6, 2)$ obtained by formula, by finding it directly, as was done for the point $(3, 4)$.
6. Find the slope of the parabola $4y = x^2$ at $(4, 4)$. Generalize this result to obtain a formula for the slope at any point.
7. Find the slope of the parabola $y = 4 - x^2$ at the point $(1, 3)$.
8. Generalize the result of Exercise 7, to obtain a formula for the slope at any point. Then find the slope at $x = 2$, $x = -2$, and $x = 0$.
9. Find the slope of the parabolic arc $y = +\sqrt{x}$ at $(\frac{1}{4}, \frac{1}{2})$. [Hint: $(\sqrt{\frac{1}{4} + h} - \frac{1}{2})/h = h/h(\sqrt{\frac{1}{4} + h} + \frac{1}{2})$.]
10. Find a formula for the slope in Exercise 9, at any point where $x > 0$.
11. Find a formula for the slope of the curve $y = 1/\sqrt{x}$, at any point where $x > 0$. [See the hint in Exercise 9.]
12. The distance, s ft, that a body falls in t sec from a position of rest, is $s = 16t^2$. Use the method of this article to find the speed at any time.
13. A stone dropped in a pool creates a circular disturbance that gradually expands. Use the method of this article to find the rate of increase in area of the disturbance, per unit increase in the radius, at the moment when the radius is 4 ft.
14. The pressure of a certain perfect gas at constant temperature is related to its volume by the equation $P = 40/V$. Use the method of this article to find the rate of change of the pressure, per unit increase in volume, at the moment when $V = 5$.

9. The Area Problem. Although the area problem does not seem, at first, to have any relation to the tangent problem, these problems are really very intimately connected, and this fact is of the utmost significance. At this point, however, we must consider the area problem by itself. Our problem is to find the area of a figure like $ABCD$ in Fig. 10, where the curve DC is the graph of a functional relation $y = f(x)$, with y positive throughout the interval considered. In geometry, the area of a plane figure is defined as the number of unit squares (i.e., 1 length unit on a side) that can be fitted to the figure so as to cover the figure completely and exactly without overlapping. We cannot find the area $ABCD$ this way, however, since squares cannot be fitted exactly to a figure with a curved boundary. Hence, we might try to approximate the area by a rectilinear figure. Thus, we might try to draw a horizontal line MN through an intermediate point of the curve, forming a rectangle as nearly as possible equal in area to that of the given figure. This requires us to estimate the intermediate point L so that the line MN will cut off one

area LNC , and add another DLM , as nearly equal as possible. The area of this rectangle = $AB \cdot KL = 5 \cdot 2 = 10$.

A more accurate approximation can be found by the following steps: draw several equally spaced vertical lines, dividing the area into vertical strips; now approximate each of these in the same way that we first approximated the whole area; and add the results. This approximation to the area is

$$\begin{aligned} \text{Area} &= AA_1 \cdot K_1L_1 + A_1A_2 \cdot K_2L_2 + A_2A_3 \cdot K_3L_3 + A_3A_4 \cdot K_4L_4 \\ &\quad + A_4B \cdot K_5L_5 \\ &= 1 \cdot \frac{3}{4} + 1 \cdot \frac{5}{4} + 1 \cdot \frac{17}{8} + 1 \cdot 3 + 1 \cdot \frac{15}{4} = 10\frac{7}{8}. \end{aligned}$$

Better approximations can be found by using more strips. But none of these approximations satisfy us as being the true area. And again, as

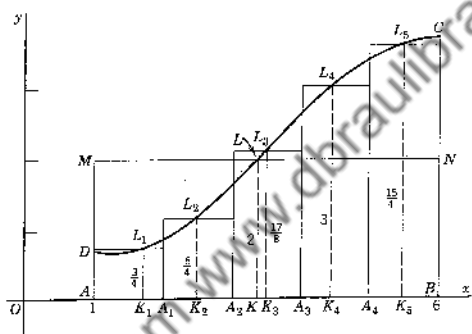


FIG. 10

with the tangent problem, we must note that the "true area" has not been defined. We feel, intuitively, the existence of such an area, however, and it is our twofold problem to formulate a satisfactory definition for it and to devise a method of calculating it.

10. The Work Problem: Generalization of the Area Problem. The problem of finding the work done by a force in moving a body a distance along a straight path is a physical analogue to the geometrical problem of area. Work done by a constant force has a simple definition, viz., work = $F \cdot s$, in which the constant force of F units moves the body a distance of s units. This definition does not provide for a varying force. What we shall call the work problem is that of formulating a definition of the work done by a variable force and of devising a method of determining it.

Let us suppose that a force of magnitude F moves a body, that we shall consider concentrated at a point and shall call a *mass-point*, along a straight line from A to B (Fig. 11); that the force has the direction of this line; and that F is a function of the distance s from a point O on this line; i.e., $F = f(s)$. If the straight line is the horizontal s -axis of a

rectangular coordinate system with origin at O , and the perpendicular axis is the F -axis, the graph of $F = f(s)$ is, in general, a curve lying above the segment AB . (We are here assuming that F is positive everywhere in this interval.) The height of each point of the graph represents the magnitude of the force associated with the point of AB that lies directly

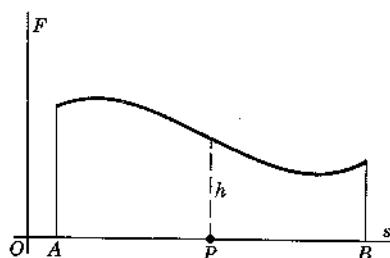


FIG. 11

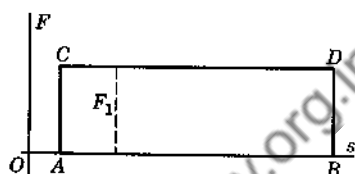


FIG. 12

below it. Thus, in Fig. 11, the length h represents the magnitude of the force at point P .

If F is a constant F_1 , the graph is a horizontal line, as in Fig. 12, and the work, $W = F_1 \cdot AB$, is equal to the number of units of area in the rectangle $ABCD$. If the force varies, suppose $\underline{F}$ and $\bar{F}$ are, respectively,

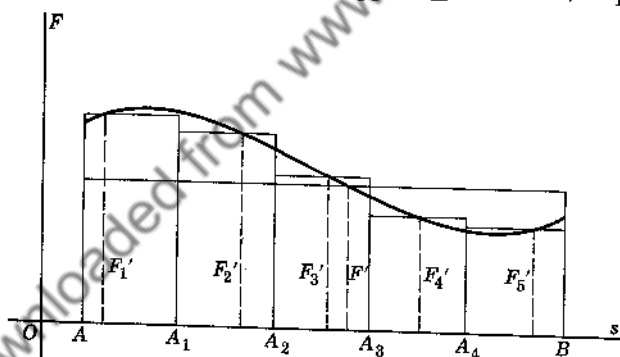


FIG. 13

the least and the greatest values F takes on in the interval from A to B . Certainly, we think of the work done by F as greater than $\underline{F} \cdot AB$ and less than $\bar{F} \cdot AB$. It is natural to approximate it by $F' \cdot AB$, where F' is some value lying between $\underline{F}$ and $\bar{F}$ (Fig. 13). As in the area problem, we would expect to reduce the error in the approximation by subdividing the distance AB into several parts, approximating the work done in moving the mass the length of each part by a product of the above form,

and adding the results. Thus we have, as a new approximation to the work done by F ,

$$F_1' \cdot AA_1 + F_2' \cdot A_1A_2 + F_3' \cdot A_2A_3 + F_4' \cdot A_3A_4 + F_5' \cdot A_4B.$$

This is precisely the method used to obtain approximations to the area under the curve. Hence, it is natural to define the magnitude of the work done by the variable force $F = f(s)$, in moving the mass-point from A to B , as the number of units in the area bounded by the graph of $F = f(s)$, the s -axis, and the ordinates erected at A and B .

The work problem is thus identified with the area problem. There are many other physical problems that can be identified with the area problem in the same way. The following are some examples: determining the mass of a rod with a variable density; finding the total pressure of a liquid on a surface, such as the side of a container; calculating the total force due to the gravitational attraction of a rod upon a unit mass; and finding the moment of inertia of a bar about a point on its axis.

There are, also, some geometric quantities, other than the area bounded by a curve, that can be found in the same way. These include lengths of curves, and volumes bounded by curved surfaces.

11. Summary. The main purpose of this chapter has been to introduce the principal problems of calculus in a way that would appeal to the intuition, without attempting to formulate them precisely. We shall see, as we proceed, that all our work centers around the tangent problem and the area problem. Although we solved the tangent problem intuitively in a few simple cases, these solutions need further scrutiny. A critical examination of them, in Chapter 2, will lead to the formulation of principles and methods of great power, which will not only make it possible to deal successfully with the tangent problem and the area problem and their many interpretations, but also will open the way to some of the deepest and most far-reaching developments in mathematics.

The Fundamental Process of Calculus

1. The Solution of the Tangent Problem Re-examined. Approximate solutions of the problems described in Chapter 1 are not hard to obtain. In some simple examples of the tangent problem we even determined the exact solution. This is not always as easy as we found it there. And, even in the simple problems discussed, it is possible that the final step to the precise solution may not have been entirely clear. Did we not take that step hastily and uncritically? Before the final step, we were surely on safe ground. There can be no doubt of our ability to obtain approximate solutions. Our main problem is that of passing from approximate solutions to the precise solution.

The procedure for obtaining the final result in the example in Art. 8, Chapter 1, was the following. (1) We combined all our approximations, and, in fact, all possible approximations, in the general expression for the slope of a chord through $P(1, 1)$ and a neighboring point of the curve, $Q(1 + h, (1 + h)^2)$, viz., $(2h + h^2)/h$. (2) Canceling h in this fraction, we obtained $2 + h$. (3) We then concluded that, since $2 + h$ could be made to approximate 2 as closely as we pleased, the slope of the tangent was 2. This conclusion, of course, involved a definition. In effect, we defined the slope of the tangent to be the number approximated by the set of numbers representing the slopes of chords. The result was then obtained by an intuitive examination of a simplified form of the slope of any chord, which was $2 + h$.

It is not suggested that there were errors in reasoning to this conclusion. It is easy, however, to make unwarranted presumptions at this point. Note that, although it is obvious that $2 + h = 2$ when $h = 0$ as well as that $2 + h$ closely approximates 2 when h is near 0, a similar statement does not apply to $(2h + h^2)/h$, which is the expression first obtained for the slope. Moreover, if h were 0, P and Q would coincide and no chord PQ would be determined. Hence, we must observe some caution in the use of $2 + h$ in place of $(2h + h^2)/h$.

When trying to obtain the number approximated by $(2h + h^2)/h$, or $2 + h$, there is a tendency to think of h as approaching 0, and of $2 + h$

approaching 2, each by a continual change, like a moving object approaching a point in its path. Since $2 + h$ is an expression for a number, this conception is not acceptable. To be sure, $2 + h$ is a variable, but a variable is merely an expression that takes on different values. No motion is involved, nor can we describe the manner by which a variable changes from one value to another except by saying that it first has one value and then another.

For a clear understanding of the process by which we can solve both the tangent problem and the area problem, we retreat to the point at which no doubt can exist, that is, to the consideration of approximations to the solution. From the approximations to the answer required we must find the answer itself. If we had only random approximations, this would be impossible. But our approximations are not random. By the process of Art. 6, Chapter I, we can obtain an ordered set of approximations, becoming better and better as we go along, and converging upon some number in such a way that the error in the approximation to this number can be made as small as we please merely by continuing the process to sufficiently many steps. Our problem then is: given a method of determining such a set of approximations, find the number approximated. The set of approximations forms a *sequence of numbers*, and it is to this concept that we turn in Art. 2.

2. Sequences. An infinite sequence of numbers is defined as follows:

DEFINITION. An infinite sequence of numbers is an ordered arrangement of infinitely many numbers, which we call the terms of the sequence, with a definite first term and a rule of formation whereby any term can be found from those that precede it.

All that is meant by "infinitely many" in this definition is that the arrangement does not terminate. No matter how many terms are written down, there are always more.

The most familiar sequence of numbers is the set of natural numbers: 1, 2, 3, $\dots$. It is useful to think of this sequence as a basic sequence, with whose terms the terms of any sequence correspond when we write them in order and number them 1st, 2nd, 3rd, etc. When the terms of a sequence are unknown we may denote them by a single letter, distinguishing between the terms, as well as indicating the order of terms, by means of subscripts 1, 2, 3, etc. Thus, the general expression for a sequence may be $a_1, a_2, a_3, \dots, a_n, \dots$. In order to make clear the rule of formation of the terms, whenever possible we express the general, or n th, term of the sequence, i.e., a_n . Thus, the n th term of the sequence of natural numbers is n itself. Further examples of sequences are the following:

$$1, 4, 9, 16, \dots, n^2, \dots, \quad (2.1)$$

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, 1/n, \dots, \quad (2.2)$$

$$0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots, (n-1)/n, \dots, \quad (2.3)$$

$$2, 4, 8, 16, \dots, 2^n, \dots, \quad (2.4)$$

$$-2, 4, -8, 16, \dots, (-2)^n, \dots, \quad (2.5)$$

$$-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, \dots, (-1)^n/n. \quad (2.6)$$

It should be observed that merely stating the n th term is sufficient to define the sequence. Thus, if the n th term is $1/n$, the 1st term is $1/1$, the 2nd term is $1/2$, etc., giving sequence (2.2). We shall often express a sequence simply by giving a formula for a_n .

If

$$a_n = (n^2 + 1)/(n^2 + 2), \quad (2.7)$$

we can write

$$a_1 = (1^2 + 1)/(1^2 + 2) = \frac{2}{3},$$

$$a_2 = (2^2 + 1)/(2^2 + 2) = \frac{5}{6},$$

$$a_3 = (3^2 + 1)/(3^2 + 2) = \frac{10}{11},$$

$$a_4 = (4^2 + 1)/(4^2 + 2) = \frac{17}{18}, \text{ etc.}$$

If

$$a_n = [n - 2 + (-1)^n]/n, \quad (2.8)$$

$$a_1 = [1 - 2 + (-1)^1]/1 = (1 - 2 - 1)/1 = -2,$$

$$a_2 = [2 - 2 + (-1)^2]/2 = (2 - 2 + 1)/2 = \frac{1}{2},$$

$$a_3 = [3 - 2 + (-1)^3]/3 = (3 - 2 - 1)/3 = 0,$$

$$a_4 = [4 - 2 + (-1)^4]/4 = (4 - 2 + 1)/4 = \frac{3}{4},$$

.....

$$a_{10} = [10 - 2 + (-1)^{10}]/10 = (10 - 2 + 1)/10 = \frac{9}{10}, \text{ etc.}$$

A superficial examination of a few sequences, like the examples given above, discloses the fact that in some sequences, as n increases, the corresponding terms increase numerically beyond all bounds. That is, no matter what number may be prescribed, we can set n large enough so that the n th term and all terms thereafter are numerically larger than this number. This is true of (2.1), (2.4), and (2.5). Such a sequence is said to *increase beyond all bounds* or *tend to infinity*. The terms of a sequence that increases beyond all bounds do not furnish better and better approximations to any number, since, whatever this number might be, eventually the terms of the sequence will be much larger.

The terms of other sequences, such as (2.2), (2.3), (2.6), (2.7), and (2.8), seem to crowd closer and closer together as we proceed. In such a case the terms of the sequence are much like the set of approximations referred to in Art. 1, from which we hope to determine the number approximated.

We shall identify the set of approximations to the solution required with a sequence of this kind, and the solution itself with what is known as the *limit* of the sequence.

EXERCISES

1. Write the values of a_5 , a_6 , and a_{100} in each of sequences (2.1) to (2.8).
2. Write the first 6 terms and the $(n + 1)$ st term of each of the following sequences:

$$(a) a_n = n + 1.$$

$$(b) a_n = (n - 1)/(n + 1).$$

$$(c) a_n = n/(n + 1).$$

$$(d) a_n = 1 - 1/n.$$

$$(e) a_n = (-1)^n n/(n + 1).$$

$$(f) a_n = 3^{n-1}.$$

$$(g) a_n = (\frac{1}{2})^n.$$

$$(h) a_n = (-1)^{n+1}/(n^2).$$

$$(i) a_n = (n^2 - 1)/(n + 1)^2.$$

$$(j) a_n = n + 1 + (-1)^n n.$$

$$(k) a_n = n[1 + (-1)^n]/(n^2 + 1).$$

$$(l) a_n = \sqrt[n]{n}.$$

3. The Limit of a Sequence. The crowding together of terms in sequences (2.2), (2.3), (2.6), (2.7), and (2.8) suggests, for each, that there may be a number, not necessarily a term of the sequence, around which the terms of the sequence cluster more and more closely as we proceed in the sequence. The further we proceed in sequence (2.2), the smaller are the terms, i.e., the closer they are to 0. In (2.3), the further we go the closer we find the terms to the number 1; e.g., if $n = 100$, $(n - 1)/n = 99/100$, differing from 1 by $1/100$, and if $n = 1000$, $(n - 1)/n = 999/1000$, differing from 1 by $1/1000$. The terms of (2.6) are obviously closer and closer to 0, the larger the value of n . The terms of (2.7) appear to approach 1 as n increases, as do those of (2.8).

The number towards which the terms of a sequence tend, as in each of these examples, is called the *limit of the sequence*. In these examples the limits were found by inspection, and that is true with a great many limits. But we cannot be satisfied with the vague description of this notion that is afforded by the preceding paragraph. The phrases "crowd together," "cluster about," "approach," and "more and more closely" are suggestive; but what do they mean, precisely? The terms of a sequence are numbers, and what we conceive to be the limit is a number. We ought to express the relation between the limit and the terms of the sequence in numerical language.

First, if L is the limit of a sequence, it appears that the terms of the sequence should differ from L by less and less as we proceed. This is true of sequence (2.3) if $L = 1$. But it would also be true if we took L to be 2. Obviously, we not only want the terms to differ from the limit by less and less as we proceed; we also want this difference eventually to become very small. Indeed, no matter how closely we may require the terms to approximate the limit, there must be terms of the sequence that meet this requirement. The terms of (2.3) approximate 1 in this way.

For, if we wish to find a term differing from 1 by only .01, we can take the 100th term, which is .99; if we wish a term differing from 1 by only .001, we can take the 1000th term, which is .999; etc. In fact, if any small positive number is assigned, by proceeding far enough in the sequence we can find a term that differs from 1 by less than this number. In order to express this requirement in a general way, we state that, if the number L is the limit of a sequence $a_1, a_2, \dots, a_n, \dots$, it must be true that, for any positive number ε (epsilon) that may be assigned, merely by proceeding sufficiently far in the sequence there can be found a term differing from L by less than ε .

This requirement is still not enough. We want to make sure not only that terms of the sequence can be found that approximate the limit within any specified tolerance, but also that a point in the sequence can be found such that all terms thereafter meet this standard of approximation. That is, we want it to be true that, when some positive ε has been assigned, we can find a value of n , say $n = m$, such that every term after the m th differs from L by less than ε . If this condition is fulfilled, the terms of the sequence, a_n , really crowd around the limit more and more closely for larger and larger values of n . Each time that we find an m th term that meets this requirement for a specified ε , this standard of approximation is met by all terms except, at most, the first m . Hence, not only are there terms very near L , but there are infinitely many that meet any standard of nearness that may be given, and only a finite number that fail to meet it.

The condition just stated is met by sequence (2.3) with the limit 1. It cannot be met by the sequence

$$0, \frac{1}{2}, -\frac{3}{4}, \frac{4}{5}, -\frac{5}{6}, \dots, (-1)^n(n-1)/n, \dots$$

Although we can find terms in this sequence that differ from 1 by as little as we please, e.g., .01, by taking n sufficiently great, there is no point in the sequence at which we can say that all terms beyond this point differ from 1 by less than .01. In fact, the negative terms differ from 1 by almost 2 when n is a large number. Moreover, this sequence does not satisfy our intuitive requirements for the possession of a definite limit, since the terms do not tend toward a single number.

The above requirements can now be formulated in a definition.

DEFINITION. The number L is said to be the limit of the sequence of numbers $a_1, a_2, \dots, a_n, \dots$, if, for each positive number ε , we can find a term in the sequence, a_m , such that for all terms following this one, i.e., for all terms a_p where $p > m$, the difference $|L - a_p|$ is less than ε .

A convenient notation for expressing the fact that a sequence has the limit L is: $\lim_{n \rightarrow \infty} a_n = L$. This is read "the limit of a_n as n increases beyond

all bounds is L ." We also express this: $a_n \rightarrow L$ as $n \rightarrow \infty$. A sequence that has a limit is said to be *convergent*, and is said to *converge*. A sequence that does not converge is called *divergent*, and is said to *diverge*. If, as n increases beyond all bounds, a_n also increases beyond all bounds, we write $a_n \rightarrow \infty$ as $n \rightarrow \infty$.

4. Remarks on Convergence. Notice that the definition just given does not require that the terms approach the limit at each step. Surely we would want to consider the sequence (2.8),

$$-2, \frac{1}{2}, 0, \frac{3}{4}, \frac{2}{5}, \frac{5}{6}, \frac{4}{7}, \frac{7}{8}, \dots, [n-2+(-1)^n]/n$$

as having the limit 1. The quantity $n-2+(-1)^n$ cannot differ from n by more than 3, and, hence, when n is large, the numerator is nearly equal to the denominator. In this sequence, however, the terms do not draw closer to 1 at every step. For example, $\frac{5}{6}$ is nearer 1 than is $\frac{2}{5}$, but $\frac{4}{7}$ is not nearer to 1 than $\frac{5}{6}$. Still, as we go on, we find that there is general progress toward 1, just as when a person takes a long step toward his goal, followed by a short step away, then a long step forward, etc.

It is to be noted that the definition of a limit does not prohibit the sequence from including the limit as one or more of its terms. The sequence $0, 1, \frac{1}{2}, 1, \frac{3}{4}, 1, \frac{7}{8}, \dots$ has the limit 1, and 1 also appears as each of the even-numbered terms. Although the definition permits this situation, in practice the limit does not usually belong to the sequence.

It is immediately apparent that many different sequences can have the same limit. For example, the sequences $a_n = 1/n$, $a_n = (-1)^n/n$, $a_n = 1/n^2$, $a_n = 1/3n$, all have the limit 0. On the other hand, a single sequence cannot have more than one limit. Let us suppose, for example, that a sequence has two different limits, A and B , differing by an amount α (alpha). Then, having set $\varepsilon = \alpha/2$, we can find a point in the sequence beyond which all terms differ from A by less than $\alpha/2$; hence, there is no point beyond which all terms differ from B by less than $\alpha/2$. That is, if the sequence has A as its limit, it cannot also have B as its limit.

Often it is easy to determine the limit of a given sequence. After reading several terms of the sequence, we see the limit intuitively. This is perhaps the most common way of finding the limit of a sequence. Intuition alone, however, is not a safe guide; results obtained intuitively should be checked to see that the definition is satisfied.

5. Theorems on Limits. If each one of a set of numbers is numerically less than some fixed number, the set of numbers is said to be *bounded*. This is the case with a convergent sequence, and we state the theorem.

THEOREM 1. The terms of a convergent sequence are bounded.

To show this we note, first, that each term is a definite number. Now, among any finite number of terms there is one or more larger than all

the rest. If L is the limit of the sequence, all but a finite number of terms, say m , differ from L by very little, e.g., by less than 1. Thus, no term can be greater in size than both $|L| + 1$ and the greatest of the first m terms.

A second theorem that is also intuitively apparent is the following.

THEOREM 2. If each of the terms of a convergent sequence is not greater than a fixed number M , then the limit is not greater than M . If each of the terms is not less than a fixed number M' , then the limit is not less than M' . That is, if $a_n \leq M$, then $\lim_{n \rightarrow \infty} a_n \leq M$; and if $a_n \geq M'$, then $\lim_{n \rightarrow \infty} a_n \geq M'$.

In order to prove the first part of this theorem, suppose that $a_n \leq M$ and $\lim_{n \rightarrow \infty} a_n = A > M$. Let $A - M = \varepsilon$. Then, for sufficiently large values of n , we must have $A - a_n < \varepsilon$. Transposing these statements, we have

$$A = M + \varepsilon \quad \text{and} \quad A < a_n + \varepsilon,$$

whence

$$a_n + \varepsilon > M + \varepsilon \quad \text{and} \quad a_n > M.$$

Since this result contradicts the hypothesis, it must follow that $\lim_{n \rightarrow \infty} a_n \leq M$.

The proof of the second part of the theorem is similar.

The four theorems that follow are indispensable for the evaluation of the limits of many sequences. They are almost obvious, and one would be likely to use them intuitively if they were not formally stated. We shall prove only one of them, to illustrate the use of the definition of a limit in the construction of a proof. These theorems apply to sequences whose terms are, respectively, the sum, difference, product, and quotient of corresponding terms of sequences with known limits.

Given two convergent sequences, a_n and b_n , with $\lim_{n \rightarrow \infty} a_n = a$ and $\lim_{n \rightarrow \infty} b_n = b$. Then

THEOREM 3. $\lim_{n \rightarrow \infty} (a_n + b_n) = a + b = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$.

THEOREM 4. $\lim_{n \rightarrow \infty} (a_n - b_n) = a - b = \lim_{n \rightarrow \infty} a_n - \lim_{n \rightarrow \infty} b_n$.

THEOREM 5. $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b = (\lim_{n \rightarrow \infty} a_n) \cdot (\lim_{n \rightarrow \infty} b_n)$.

THEOREM 6. $\lim_{n \rightarrow \infty} (a_n/b_n) = a/b = (\lim_{n \rightarrow \infty} a_n)/(\lim_{n \rightarrow \infty} b_n)$, provided that $b_n \neq 0$ for sufficiently large n , and $b \neq 0$.

PROOF OF THEOREM 5. Since $\lim_{n \rightarrow \infty} a_n = a$ and $\lim_{n \rightarrow \infty} b_n = b$, we know, by the definition of a limit, that we can make the difference $|a_n - a|$ and

$|b_n - b|$ as small as we please by taking n sufficiently great. That is, for any positive number ε that might be chosen, we can find an integer m such that

$$|a_p - a| < \varepsilon \quad \text{and} \quad |b_p - b| < \varepsilon \quad \text{if } p > m.$$

Now, form the difference $a_p \cdot b_p - ab$. This can be written $a_p(b_p - b) + b(a_p - a)$. Since the sequence a_n is bounded, we can put $|a_p|$ less than some positive number M . Hence

$$\begin{aligned} |a_p b_p - ab| &= |a_p(b_p - b) + b(a_p - a)| \\ &\leq M|b_p - b| + |b| |a_p - a| \\ &< M\varepsilon + |b|\varepsilon \\ &< (M + |b|)\varepsilon. \end{aligned}$$

Since M and b are fixed numbers, by choosing ε small enough we can make $|a_p b_p - ab|$ as small as we please, by taking p sufficiently great. That is, $\lim_{n \rightarrow \infty} a_n b_n = ab$. (If we insist upon making $|a_p b_p - ab| < \varepsilon'$, all we need do is choose $\varepsilon < \varepsilon'/(M + |b|)$.)

6. Limits of Some Important Sequences. It is a fundamental assumption concerning the system of natural numbers that, given any number N , no matter how large, there are natural numbers larger than N , i.e., that the terms of the sequence $1, 2, 3, \dots, n, \dots$ really increase beyond all bounds. From this assumption it is easy to prove that $\lim_{n \rightarrow \infty} (1/n) = 0$.

This fact is, of course, familiar and appears quite obvious. In order to prove it by reference to the definition of a limit, imagine any positive ε that might be chosen. In accordance with the above assumption, we can take $m > 1/\varepsilon$. Then $1/m < \varepsilon$, by the rules of inequalities. Also, if $p > m$, then $1/p < 1/m < \varepsilon$.

The knowledge that $\lim_{n \rightarrow \infty} (1/n) = 0$ is very helpful in dealing with limits of more complicated sequences. For example, we see that

$$\begin{aligned} \lim_{n \rightarrow \infty} (n+1)/n &= \lim_{n \rightarrow \infty} (1 + 1/n) \\ &= \lim_{n \rightarrow \infty} 1 + \lim_{n \rightarrow \infty} (1/n), \text{ by Theorem 3, Art. 5} \\ &= 1. \end{aligned}$$

(Lim 1 is clearly 1, every term of the sequence being equal to 1.)

Similarly,

$$\begin{aligned}
 \lim_{n \rightarrow \infty} (n + a)/n &= \lim_{n \rightarrow \infty} (1 + a \cdot [1/n]) \\
 &= \lim_{n \rightarrow \infty} 1 + \lim_{n \rightarrow \infty} a [1/n], \text{ by Theorem 3, Art. 5} \\
 &= 1 + (\lim_{n \rightarrow \infty} a) \cdot (\lim_{n \rightarrow \infty} [1/n]), \text{ by Theorem 5, Art. 5} \\
 &= 1 + a \cdot 0 = 1.
 \end{aligned}$$

A more general form of the fundamental assumption concerning natural numbers stated above, known as the Theorem of Eudoxus, is the following: given any two positive numbers P and Q , it is possible to choose the integer n so that $nP > Q$. This principle reduces to the assumption stated at the beginning of this article if $P = 1$. Using this theorem, we can prove that $b^n \rightarrow \infty$ as $n \rightarrow \infty$, if $b > 1$. For, if $b > 1$, we can write $b = 1 + \alpha$, where $\alpha > 0$. Then, using the binomial theorem,

$$\begin{aligned}
 b^n &= (1 + \alpha)^n = 1 + n\alpha + \frac{n(n-1)}{2} \alpha^2 + \dots \\
 &> 1 + n\alpha, \text{ if } n > 1.
 \end{aligned}$$

Now, $n\alpha \rightarrow \infty$ as $n \rightarrow \infty$, since, for any positive number Q , we can make $n\alpha > Q$. Therefore $(1 + n\alpha) \rightarrow \infty$, and hence $b^n \rightarrow \infty$ as $n \rightarrow \infty$.

If $b < 1$, then $1/b > 1$, and $(1/b)^n \rightarrow \infty$ as $n \rightarrow \infty$. Thus, for any $\varepsilon > 0$, we can choose m so that $(1/b)^n > 1/\varepsilon$, if $n > m$. That is, $1/b^n > 1/\varepsilon$ and $b^n < \varepsilon$. Hence, $b^n \rightarrow 0$ as $n \rightarrow \infty$.

Of course, if $b = 1$, the sequence $b^n \rightarrow 1$ as $n \rightarrow \infty$, since every term is equal to 1.

Similarly, if $a =$ a constant different from 0, as $n \rightarrow \infty$,

$$ab^n \rightarrow \infty \text{ if } b > 1 \quad \text{and} \quad ab^n \rightarrow 0 \text{ if } b < 1.$$

Thus, the terms of a geometrical progression tend to ∞ if the common ratio > 1 , and to 0 if the common ratio < 1 .

EXAMPLE 1. Find $\lim_{n \rightarrow \infty} \frac{2n + 4}{n - 2}$.

SOLUTION. Since n cannot be 0, we can divide numerator and denominator by n , obtaining

$$\frac{2n + 4}{n - 2} = \frac{2 + 4/n}{1 - 2/n}.$$

Hence, by Theorem 6, Art. 5,

$$\lim_{n \rightarrow \infty} \frac{2n+4}{n-2} = \frac{\lim (2+4/n)}{\lim (1-2/n)} = \frac{2}{1} = 2.$$

EXAMPLE 2. Prove that $(1/n^n) \rightarrow 0$ as $n \rightarrow \infty$.

SOLUTION. We can write $1/n^n = (1/n)(1/n^{n-1})$. Now, for $n = 2, 3, \dots$, we have $(1/n^{n-1}) < 1$. Hence

$$1/n^n = (1/n)(1/n^{n-1}) < 1/n.$$

Since $1/n \rightarrow 0$ as $n \rightarrow \infty$, it follows that a positive quantity less than $1/n$ must tend to 0, also. Therefore, $1/n^n \rightarrow 0$ as $n \rightarrow \infty$.

EXERCISES

1. Prove that each of the following sequences converges to 0.

(a) $a_n = 1/2n$.

(b) $a_n = 1/n^2$.

(c) $a_n = (n+1)/n^2$.

(d) $a_n = n/(n+1)^2$.

(e) $a_n = 1/b^n$, $b > 1$.

(f) $a_n = b^n/c^n$, $b < c$.

(g) $a_n = 10^{10}/n$.

(h) $a_n = 1/(2^n + n)$.

2. Prove that $1/\sqrt[n]{n} \rightarrow 0$ as $n \rightarrow \infty$. [Hint: Use an indirect proof: If this is not the case, then $1/n$ does not tend to 0 as $n \rightarrow \infty$.]

3. Use the method of Exercise 2 to show that $1/n^p \rightarrow 0$ as $n \rightarrow \infty$, if p is a positive fraction.

4. Find the limit of each of the following sequences, and prove that you are right.

(a) $a_n = (2n+1)/n$.

(b) $a_n = 1/n - b^n$, $0 < b < 1$.

(c) $a_n = (n^2 - 1)/(n^2 + 2n + 1)$. [Hint: Divide numerator and denominator by n^2 .]

(d) $a_n = (n^2 + n + 1)/(3n^2 - 1)$.

(e) $a_n = (n^3 - n + 3)/(n^3 + 1)$.

(f) $a_n = (n^3 + 3n^2 + 1)/(n^4 - 2n + 3)$.

(g) $a_n = (2n+1)(1/2)^n/n$.

(h) $a_n = 1/n^2 + (n+1)/n - (n^2-1)/(n^2+n+1)$.

5. What can be said about the limit of

$$\frac{c_0 n^k + c_1 n^{k-1} + \dots + c_k}{b_0 n^h + b_1 n^{h-1} + \dots + b_h} \text{ as } n \rightarrow \infty,$$

if both k and h are positive integers and (a) $k > h$? (b) $k < h$? (c) $k = h$?

6. What can be said about the limit of

$$\frac{\sqrt[k]{kn+1}}{\sqrt[h]{hn+1}} \text{ as } n \rightarrow \infty,$$

if r and s are both positive integers and (a) $r > s$? (b) $r < s$? (c) $r = s$?

7. Prove the following corollary to Theorem 2, Art. 5. If the sequences a_n and b_n both converge, and if $a_n \leq b_n$ for all values of n , then $\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n$.

7. Solution of the Tangent Problem. Let us reconsider the slope problem in the light of the discussion of sequences. Again we shall try to find the slope of the parabola $y = x^2$ at $P(1, 1)$. The student should again read Art. 6, Chapter 1.

As in that discussion, we draw chords through P and Q_1, Q_2 , etc., but with the abscissas of Q_1, Q_2 , forming a sequence having the limit 1 (Fig. 1). If the sequence of abscissas is

$$1 + 1, 1 + 1/2, 1 + 1/3, \dots, 1 + 1/n, \dots,$$

the slopes of the corresponding chords also form a sequence, namely:

$$a_1 = \text{slope of } PQ_1 = (4 - 1)/(2 - 1) = 3,$$

$$\begin{aligned} a_2 = \text{slope of } PQ_2 &= [(1 + 1/2)^2 - 1]/[(1 + 1/2) - 1] \\ &= (1 + 1/4)/(1/2) = 2 + 1/2, \end{aligned}$$

$$\begin{aligned} a_n = \text{slope of } PQ_n &= [(1 + 1/n)^2 - 1]/[(1 + 1/n) - 1] \\ &= (2/n + 1/n^2)/(1/n) = 2 + 1/n, \end{aligned}$$

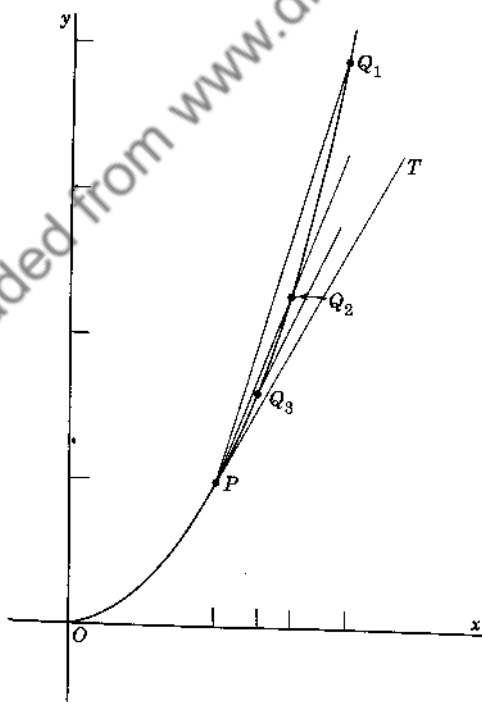


FIG. 1

Since it is evident that the sequence of slopes a_n corresponds to a sequence of chords PQ_n which approximate our intuitive idea of the tangent at P better and better as n increases, we *define* the tangent PT as the line through P with slope $= \lim_{n \rightarrow \infty} a_n$. That is, slope $PT = 2$, and the equation of PT is

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

Neither of the questions that plagued us in connection with the intuitive solution of this problem in Chapter 1 casts any shadow on this solution. There is no question of $1/n$ becoming 0, since there is no value of n that makes this possible. In addition, the approach of $2 + 1/n$ to 2 is strictly a numerical one through a sequence of numerical values.

Some practical obstacles are still to be surmounted. For all but the simplest functions, the formation of a sequence of approximating slopes will be laborious. Also, there is an arbitrariness in the choice of a particular approximating sequence. How shall we decide what values of abscissas of Q_n to use? If different sequences are employed, can we be assured of arriving at the same value for the slope of PT ? We answer these questions by generalizing the sequence of abscissas of Q_n . Let Q have the abscissa $1 + h$, and the corresponding ordinate $(1 + h)^2$. The slope of PQ is then $(2h + h^2)/h$, or $2 + h$. This seems to bring us back to the earlier, rather vague, treatment of Art. 8, Chapter 1. In the light of the study of sequences, however, we now require that h take on a sequence of values tending to the limit 0, but never being equal to 0. Since each value of h yields a corresponding value of $(2h + h^2)/h$, or $2 + h$, to which this fraction is always equal when $h \neq 0$, a sequence of values of $2 + h$ is thus created. Obviously, any sequence of values of h that has the limit 0 must yield a sequence of values of $2 + h$ that has 2 for its limit. Hence, there is no problem as to the choice of a particular sequence, since all lead to the same result.

The replacement of a particular approximating sequence by a general approximating function of h , like $(2h + h^2)/h$, makes the formation of a particular sequence unnecessary. The advantage gained by this change is immeasurable. Indeed, if we were required to write the terms of an approximating sequence for every limit of the sort we are here concerned with, the tedium and awkwardness of the manipulations of calculus would be intolerable. In order to exploit this advantage, however, we must gain a clear conception of the idea of a limit approached by a function of a variable like h , which, unlike the integer n , can assume any value in an interval.

8. Limit of a Function. The limit of a sequence is a special case of the *limit of a function*: it is the limit of a function of the integral variable n

as n increases beyond all bounds. An integral variable is one that takes on only integral values. We now try to find the limit of $f(x)$ as x approaches some fixed number, i.e., as x takes on a sequence of values tending to a definite limit a , a process which we denote by $x \rightarrow a$. In formulating this concept we keep in mind the question that brought it to our attention. We want it to apply to such functions as $f(h) = (2h + h^2)/h$ as $h \rightarrow 0$. Consequently, we do not allow h to be 0. Moreover, we want to permit h to assume the values in any sequence tending to 0 (not containing 0 as a term), and we want the sequence of corresponding values of $f(h)$ to have the same limit in all cases. In saying that $f(h) \rightarrow 2$ as $h \rightarrow 0$ we mean that $f(h)$ approximates 2 if h approximates 0, and, indeed, that the difference $|f(h) - 2|$ can be made as small as we wish by making $|h|$ small enough. Putting this idea into precise language, we say that, for any ε one might choose, it must be possible to find another positive number δ (delta) such that $|f(h) - 2| < \varepsilon$ if $0 < |h| < \delta$. Generalizing, we define $\lim_{x \rightarrow a} f(x)$, read "limit of $f(x)$ as x tends to a ," as follows.

DEFINITION. The number b is said to be the limit of $f(x)$ as x tends to a if, for any positive number ε that may be chosen, a second positive number δ can be found, such that $|f(x) - b| < \varepsilon$ if $0 < |x - a| < \delta$.

The restriction, $0 < |x - a|$, in this definition, reflects the idea that $f(x)$ must be near b when x is near a . What may happen to $f(x)$ for $x = a$, whatever its value may be or whether it has a value, is irrelevant to the question of the limit to which $f(x)$ tends as x tends to a . The definition contains this restriction for the express purpose of making it apply to such functions as $(2h + h^2)/h$ as $h \rightarrow 0$. Moreover, because of it, we can replace $(2h + h^2)/h$ by $2 + h$, which is dealt with more easily (see Art. 12).

The relationship of the idea of the limit of a sequence to that of the limit of a function is brought out by the following observations. First, if $\lim_{x \rightarrow a} f(x) = b$, then the values of x in every sequence having a as its limit, and not including a among its terms, must yield a sequence of values of $f(x)$ whose limit is b . Hence, the limit b could be found by means of any one of these sequences of x values. Second, even though we seldom need to express such a sequence in detail, its existence, together with a corresponding sequence for $f(x)$, having the limit b , is the basic concept on which the idea of the limit of a function rests.

Instead of the statement, $\lim_{x \rightarrow a} f(x) = b$, it is often advantageous to express the same fact by saying that

$$f(x) = b + \varepsilon(x) \quad \text{wherein} \quad \lim_{x \rightarrow a} \varepsilon(x) = 0.$$

The notation $\varepsilon(x)$ indicates that ε is dependent on x . It is easily seen that these statements are equivalent.

9. Examples. Theorems on Limits. Limits of many simple functions offer no difficulties. In fact, it will often appear that the language of limits is overelaborate and unnecessary in connection with many familiar functions. This is true in Example 1. Examples 2-6 are less simple, and the limits are much less obvious.

EXAMPLE 1. Let $f(x) = 3x$, and let it be required to find $\lim_{x \rightarrow 2} f(x)$. Clearly, $3x$ tends to 6 as x tends to 2. The fact that $3x = 6$ when $x = 2$ is interesting, but irrelevant in connection with $\lim_{x \rightarrow 2} 3x$.

EXAMPLE 2. Let $f(x) = 1/(1-x)$ for $x \neq 1$, [$f(x)$ is not defined for $x = 1$], and let us try to find $\lim_{x \rightarrow 1} f(x)$. We see that any sequence of values of x that tends to 1 corresponds to a sequence of values of $1/(1-x)$ that increases beyond all bounds. Hence, there is no number to which $f(x)$ tends as x tends to 1. Although $\lim_{x \rightarrow 1} 1/(1-x)$ does not exist, we often say that $1/(1-x)$ tends to ∞ as x tends to 1.

EXAMPLE 3. Let $f(x) = (x-1)/(x^2-1)$ for $x^2 \neq 1$ [$f(x)$ is not defined for $x = 1$ or $x = -1$]. We want to find $\lim_{x \rightarrow 1} f(x)$ and $\lim_{x \rightarrow -1} f(x)$, if these limits exist. First, consider $\lim_{x \rightarrow 1} (x-1)/(x^2-1)$. Since, by the definition of a limit, we are not required to consider $x = 1$ and really have no right to consider it, we can cancel the factor $x-1$ from both numerator and denominator, obtaining $1/(x+1)$. If x could be equal to 1, this cancellation would not be permissible. We can now write $\lim_{x \rightarrow 1} (x-1)/(x^2-1) = \lim_{x \rightarrow 1} 1/(x+1)$. It is easy to see that, if x takes on any sequence of values tending to 1, then $1/(x+1)$ tends to $\frac{1}{2}$. Hence, $\lim_{x \rightarrow 1} (x-1)/(x^2-1) = \frac{1}{2}$. The function is not defined for $x = 1$, and thus we cannot say that $f(1) = \frac{1}{2}$, even though $\lim_{x \rightarrow 1} f(x) = \frac{1}{2}$.

Writing $(x-1)/(x^2-1) = (x-1)/(x-1)(x+1)$, we see that, as x takes on values near -1 , this fraction increases numerically beyond all bounds. The cancellation of $x-1$ is allowable here also, since we are interested in values of x near -1 and these need not include $x = 1$. This function is identical with $1/(x+1)$ except for the value $x = 1$, which belongs to the range of the latter function but not to the given one. The graph is the rectangular hyperbola shown in Fig. 2, with the single point $(1, \frac{1}{2})$ deleted. The meaning of the ε, δ requirement of the definition of the limit is illustrated at this point. Thus, the graphical interpretation of the statement $\lim_{x \rightarrow 1} f(x) = \frac{1}{2}$ is the following: if two horizontal lines are

drawn, one ϵ units above point $P(1, \frac{1}{2})$ and the other ϵ units below P , we can now draw two vertical lines, one δ units to the right of P and the other δ units to the left of P , so that the part of the graph that lies between

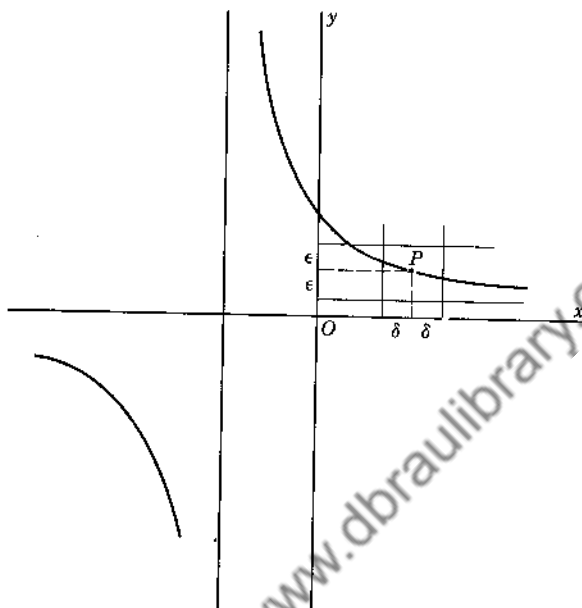


FIG. 2

the verticals must also lie between the horizontal. P itself is not a part of the graph, since $f(x)$ is not defined for $x = 1$.

EXAMPLE 4. Let $f(x) = 1$ for all integral values of x , and $f(x) = \frac{1}{2}$ for all other real values of x , and let it be required to find $\lim_{x \rightarrow 1} f(x)$. The

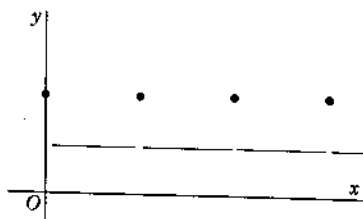


FIG. 3

graph of this function is shown in Fig. 3, and it is readily seen that $\lim_{x \rightarrow 1} f(x) = \frac{1}{2}$. The value of $f(1)$ is not $\frac{1}{2}$, however, but 1.

EXAMPLE 5. Let $f(x) = 0$ for $x = 0$, $f(x) = 3$ for $0 < x \leq 1$, and, in general, $f(x) = 3n$ for $(n - 1) < x \leq n$, where n is any positive integer.

(This is the function that tells the number of cents postage needed for x ounces of first-class mail.) The graph is shown in Fig. 4. We see that $f(x)$ fails to have a limit as $x \rightarrow n$. For, no matter how close we draw vertical lines on opposite sides of any integral value of x , there are always values of $f(x)$ between these verticals that differ by 3. Therefore, the definition of the limit of $f(x)$ cannot be satisfied by any value of L if a is an integer.

EXAMPLE 6. Let $f(x) = (\log_{10} x)/(x - 1)$ for $x \neq 1$, and let us try to find $\lim_{x \rightarrow 1} f(x)$. Clearly, the denominator tends to 0 as x tends to 1. Our

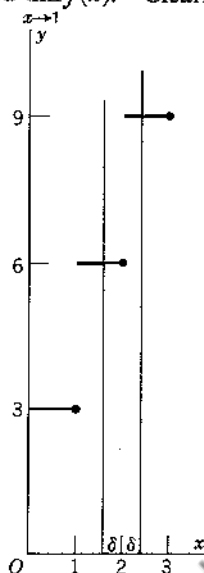


FIG. 4

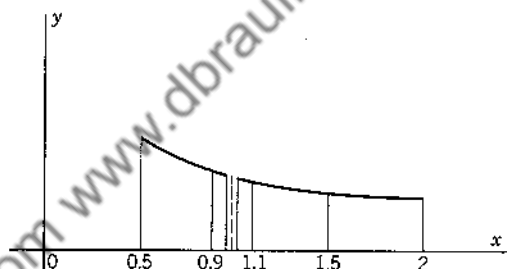


FIG. 5

knowledge of the logarithms of numbers tells us that $\log_{10} x$ also tends to 0 as x tends to 1. But, unlike Example 3, in which both numerator and denominator tend to 0 because of the common factor $x - 1$, which we can cancel, there is no function by which we can replace $(\log_{10} x)/(x - 1)$ so as to see intuitively what limit is approached. Some evidence is afforded by the graph in Fig. 5, drawn from the following table.

x	2	1.5	1.1	1.01	.99	.9	.5
$f(x)$	.3	.35	.41	.43	.44	.46	.61

From this it appears that $\lim_{x \rightarrow 1} (\log_{10} x)/(x - 1)$ is approximately .435.

Theorems 3-6 of Art. 5, on sequences, hold equally well for sums, differences, products, and quotients of functions. We state them without proof.

Given $\lim_{x \rightarrow a} f(x) = b$ and $\lim_{x \rightarrow a} g(x) = c$, then

THEOREM 1. $\lim_{x \rightarrow a} [f(x) + g(x)] = b + c$;

THEOREM 2. $\lim_{x \rightarrow a} [f(x) - g(x)] = b - c$;

THEOREM 3. $\lim_{x \rightarrow a} [f(x) \cdot g(x)] = b \cdot c$;

THEOREM 4. $\lim_{x \rightarrow a} [f(x)/g(x)] = b/c$, provided that $c \neq 0$.

EXERCISES

1. Find the following limits, stating a theorem to justify each step:

(a) $\lim_{x \rightarrow 2} 5x$.

(b) $\lim_{x \rightarrow 3} x(x + 2)$.

(c) $\lim_{x \rightarrow 2} (x^2 + 3x + 2)$.

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

2. Show that $\lim_{x \rightarrow 1} \frac{x^3 - 1}{x - 1} = 3$.

3. Show that $\lim_{x \rightarrow -2} \frac{x^2 - 4}{x + 2} = -4$.

4. Show that $\lim_{x \rightarrow 2} \frac{x - 2}{x^2 - 4} = \frac{1}{4}$.

5. What can be said about $\lim_{x \rightarrow 0} 1/x$? About $\lim_{x \rightarrow 2} 2/(x - 2)$?

6. Prove Theorem 1. [Hint: Write $f(x) = b + \varepsilon(x)$, $g(x) = c + \varepsilon'(x)$, in which $\varepsilon(x)$ and $\varepsilon'(x)$ tend to 0 with x .]

7. Prove Theorem 3.

8. Estimate $\lim_{x \rightarrow 0} \frac{\sin x}{x}$ by means of a graph ($x =$ a number of radians). See Exercise 8(d), Art. 4, Chapter 1.

9. Estimate $\lim_{x \rightarrow 0} \frac{10^x - 1}{x}$. See Exercise 8(c), Art. 4, Chapter 1.

10. Estimate $\lim_{x \rightarrow 0} x \log x$. See Exercise 8(a), Art. 4, Chapter 1.

10. Continuous Functions. The definition of the limit of $f(x)$ as x tends to a makes no use of $f(a)$, nor does it require $f(a)$ to exist. This restriction has the express purpose of making the definition apply to functions that fail to have meaning for $x = a$ but that have definite values for all values of x near a . This was the condition that applied to the function of h representing the slope of a chord approximating the tangent. In general, therefore, we cannot find $\lim_{x \rightarrow a} f(x)$ by substituting $x = a$ in $f(x)$.

There is a large class of functions, however, in which the limits can be found by this substitution process, and when we know that we are dealing with such a function it is permissible, and usually advisable, to find them this way. Such functions are called *continuous functions*. A function is said to be continuous at a point $x = a$ if it satisfies the following definition.

DEFINITION. The function $f(x)$ is continuous at $x = a$ if (1) $f(x)$ tends to a limit as x tends to a , (2) $f(x)$ has a value at $x = a$, and (3) the limit as $x \rightarrow a$ is the same as the value of the function at $x = a$. More briefly, $f(x)$ is continuous at $x = a$ if $\lim_{x \rightarrow a} f(x) = f(a)$.

This definition refers to continuity at a point, i.e., at a particular value of x . Continuity in an interval is defined as follows.

DEFINITION. The function $f(x)$ is continuous in an interval if it is continuous at every point in the interval.

Graphically, we think of a function continuous in an interval as one represented by a curve in which there are no breaks. A function that is not continuous is called *discontinuous*. Among the examples in Art. 9, the function in Example 1 is continuous everywhere. That in Example 2 is discontinuous at $x = 1$, since it is not defined there nor does $f(x)$ tend to a limit as $x \rightarrow 1$. The function in Example 3 is discontinuous at $x = 1$, and $x = -1$, since it is not defined for those values. In addition, as $x \rightarrow -1$ this function does not tend to a limit. The function in Example 4 is discontinuous at each integral value of x , since the limit approached as $x \rightarrow$ an integer is different from the functional value at that point. The function in Example 5 is discontinuous at each integral value of x . In this example the function has no definite limit as $x \rightarrow$ an integer. In Example 6 the function is discontinuous at $x = 1$, since it is not defined there.

It is natural to try to obtain $\lim_{x \rightarrow a} f(x)$ by substituting $x = a$ in $f(x)$.

The fact that this is possible only if the function is continuous makes the continuity properties of a given function an important item in the study of the function. The continuity of many algebraic functions is discussed in Art. 12. The results found there follow from the theorems stated in Art. 11.

11. Theorems on Continuous Functions. The four theorems on limits of functions, stated in Art. 9, lead to four similar theorems on continuous functions, which may be stated as follows. If $f(x)$ and $g(x)$ are both continuous at $x = a$, then

THEOREM 1. The sum $f(x) + g(x)$ is continuous at $x = a$;

THEOREM 2. The difference $f(x) - g(x)$ is continuous at $x = a$;

THEOREM 3. The product $f(x) \cdot g(x)$ is continuous at $x = a$;

THEOREM 4. The quotient $f(x)/g(x)$ is continuous at $x = a$, if $g(a) \neq 0$.

Three other theorems that are easily grasped from graphical considerations are the following.

THEOREM 5. If $f(x)$ is continuous in the closed interval $a \leq x \leq b$, then $f(x)$ takes on every value between $f(a)$ and $f(b)$ for at least one value of x in the interval. Thus, in Fig. 6, α lies between $f(a)$ and $f(b)$, and

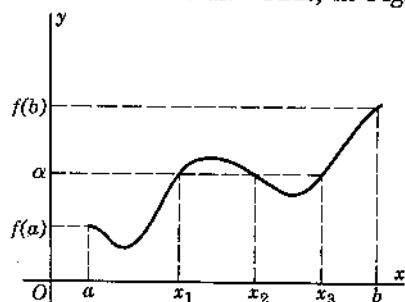


FIG. 6

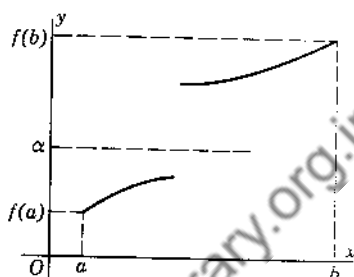


FIG. 7

$f(x) = \alpha$ for $x = x_1, x_2$, and x_3 . This statement may fail to hold if $f(x)$ is discontinuous at some point in the interval (Fig. 7).

THEOREM 6. If $f(x)$ is continuous in the closed interval $a \leq x \leq b$, then $f(x)$ takes on a maximum value and a minimum value within the interval. This theorem is illustrated in Fig. 8, in which the minimum

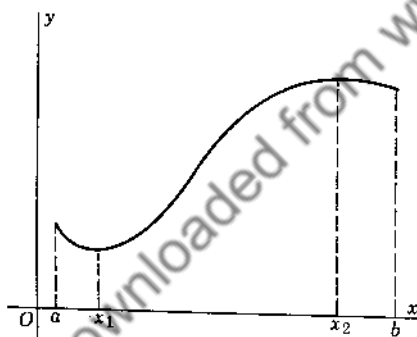


FIG. 8

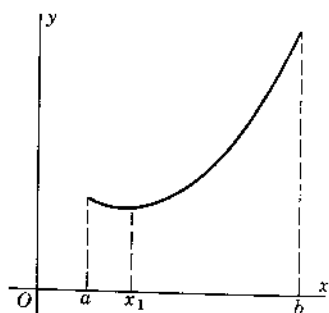


FIG. 9

occurs at $x = x_1$ and the maximum occurs at $x = x_2$. In Fig. 9, the minimum occurs at $x = x_1$ and the maximum occurs at the end of the interval $x = b$. In Fig. 10, the function is discontinuous at $x = c$, where the curve has a vertical asymptote, and there is no maximum in the interval.

THEOREM 7. If $f(x)$ is continuous at a point $x = c$ in the interior of the interval in which $f(x)$ is defined, and if $f(c)$ is greater than some number α , then there is an interval including c throughout which $f(x) > \alpha$. Thus, in Fig. 11, $f(c) > 0$, and, because $f(x)$ is continuous, it cannot

suddenly fall below 0, but must remain greater than 0 throughout an interval such as that between the dotted lines. If $f(x)$ is continuous at a point $x = c$ in the interior of the interval in which $f(x)$ is defined, and if $f(c)$ is less than some number β , then there is an interval including c throughout which $f(x) < \beta$.

12. Continuity of Algebraic Functions. Many of the functions important in applications of calculus are continuous. The continuity of many common algebraic functions follows from Theorems 1-4 of Art. 11.

Since a polynomial is formed from sums and products of constants and integral powers of x , every polynomial is continuous for all real values of x . The quotient of two polynomials, i.e., a rational function, is continuous for all values except those which reduce the denominator to

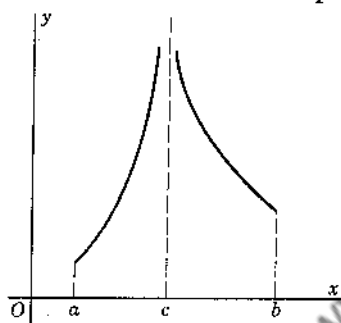


FIG. 10

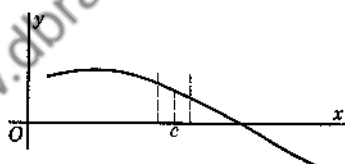


FIG. 11

0. For example, (1) $5x^3 - 6x^2 + 2x + 10$ is continuous for all real values of x ; (2) $(7x^2 + 1)/(x - 2)$ is continuous for all real values of x except $x = 2$; (3) $(x^2 + x - 6)/(x^2 - 4)$ is continuous for all real values of x except $x = 2$ and $x = -2$.

The quotient $(7x^2 + 1)/(x - 2)$ in Example 2 is discontinuous at $x = 2$, since no value is defined at $x = 2$, and, moreover, the quotient does not tend to a limit as x tends to 2. The situation is not quite the same with the quotient $(x^2 + x - 6)/(x^2 - 4)$ at $x = 2$ in Example 3. This fraction does not define a value at $x = 2$, but it does approach a limit as x approaches 2. By the method used in Example 3 of Art. 9, we see that this limit is the same as that of $(x + 3)/(x + 2)$, which is obtained by writing $(x^2 + x - 6)/(x^2 - 4) = [(x - 2)(x + 3)]/[(x - 2)(x + 2)]$ and canceling the factor $x - 2$. Since the given fraction is equal to $(x + 3)/(x + 2)$ for all values except $x = 2$, its limit as $x \rightarrow 2$, upon which the value $x = 2$ has no bearing, is the same as that of $(x + 3)/(x + 2)$. But the latter fraction is continuous at $x = 2$, and hence this limit is $(2 + 3)/(2 + 2)$, or $5/4$. Since the fraction $(x^2 + x - 6)/(x^2 - 4)$ does not define a value at $x = 2$, the function defined by this fraction alone is

discontinuous at $x = 2$. But if we add to the definition of the function the statement that the function is equal to $5/4$ at $x = 2$, we shall have a function which is continuous at $x = 2$. The function is still discontinuous at $x = -2$, and it cannot be made continuous by any way of defining it at $x = -2$, since the fraction $(x^2 + x - 6)/(x^2 - 4)$ fails to approach a limit as $x \rightarrow -2$.

The student should notice particularly that the procedure above, whereby $\lim_{x \rightarrow 2} \frac{(x-2)(x+3)}{(x-2)(x+2)}$ was found by canceling the factor $x - 2$, is a method of finding the limit of the fraction as x approaches 2. It does not determine the value of $f(2)$, for there is no value of $f(2)$. This method of finding limits is very common. It consists in replacing a function that is not continuous at the point in question by another function that has the same limit and is continuous, and in which we can therefore find the limit by simply inserting the value of x that is under consideration. Thus, in Art. 7, the limit of $(2h + h^2)/h$ was identified with that of $2 + h$. The student may, at first, feel that this is circular reasoning, for we first assert that h cannot be equal to 0 and hence we can divide by h , and then we finish by putting $h = 0$ in the quotient $2 + h$. It should be observed that $(2h + h^2)/h$ and $2 + h$ are not identically equal, but are equal for all values of h that are relevant in determining the limit, i.e., all values except 0. They must therefore have the same limit. The second function, however, is continuous, and its limit can be found by putting $h = 0$. This could not be done in the fraction.

We close this discussion of continuous functions with the following general theorem.

If $u = f(x)$ is a continuous function of x at $x = a$, $u = b$, and if $y = \phi(u)$ is a continuous function of u at $u = b$, $y = c$, then $y = \phi[f(x)] = F(x)$ is a continuous function of x at $x = a$, $y = c$. This theorem follows from the facts that $u \rightarrow b$ as $x \rightarrow a$ and $y \rightarrow c$ as $u \rightarrow b$, and $b = f(a)$ and $c = \phi(b) = \phi[f(a)] = F(a)$.

For continuity in an interval, this theorem becomes: If $u = f(x)$ is continuous in the interval from $x = a$ to $x = b$, and if $y = \phi(u)$ is continuous in the interval between the minimum and maximum values assumed by $f(x)$ in the interval from a to b , then $y = \phi[f(x)] = F(x)$ is continuous in the interval from $x = a$ to $x = b$.

EXERCISES

1. In what range is each of the following functions continuous?

(a) $f(x) = x^2$.

(b) $f(x) = x^2 - 3x + 2$.

(c) $f(x) = \frac{1}{x-1}$.

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

$$(e) f(x) = x(x^3 + 4).$$

$$(f) f(x) = \frac{1}{x} + \frac{3}{x-1}.$$

$$(g) f(x) = \frac{x^2 - 1}{x - 1}.$$

$$(h) f(x) = \frac{x^2 - 9}{x^2 - 2x - 3}.$$

$$(i) f(x) = \frac{x^2 - x}{x - 1} \text{ for } x \neq 1, f(x) = 1 \text{ for } x = 1.$$

$$(j) f(x) = \frac{x^2 + 2x - 3}{x^2 - 1} \text{ for } |x| \neq 1, f(1) = 2, f(-1) = 1.$$

$$(k) f(x) = 3u^2 - 1, u = \frac{1}{x}.$$

$$(l) f(x) = 2u + \frac{1}{u}, u = x - 3.$$

2. Find the following limits.

$$(a) \lim_{x \rightarrow 2} 8x + 4;$$

$$(b) \lim_{x \rightarrow 1} \frac{x^2 - x}{x - 1};$$

$$(c) \lim_{x \rightarrow -1} \frac{x^2 + 2x + 1}{x + 1};$$

$$(d) \lim_{x \rightarrow 5} \frac{2x}{x^2 - 25};$$

$$(e) \lim_{x \rightarrow 0} \frac{5 + 3x + x^2}{1 - x};$$

$$(f) \lim_{x \rightarrow -3} \frac{x^3 + 27}{x^2 + 5x + 6};$$

$$(g) \lim_{x \rightarrow 1} 9x^{10} - 7x^6 + 11x^5 + 4x^3 - 20x^2 + x - 8.$$

13. The Tangent Problem. Recapitulation. We now have all the concepts necessary for the solution of the tangent problem, completely, rigorously, and expeditiously. The slope of the tangent to the curve $y = x^2$ at the point $(1, 1)$ is simply

$$\lim_{h \rightarrow 0} \frac{(1+h)^2 - 1}{h} = \lim_{h \rightarrow 0} \frac{2h + h^2}{h} = \lim_{h \rightarrow 0} (2 + h) = 2.$$

The steps in this process were discovered intuitively in Chapter 1. They have now been made clear by the discussion of limits of sequences and of functions.

Similarly, the slope of the tangent to this curve at $(\frac{1}{2}, \frac{1}{4})$ is found by taking the limit of a function representing the slope of a chord through this point and any neighboring point $[\frac{1}{2} + h, (\frac{1}{2} + h)^2]$ (Fig. 12).

$$\lim_{h \rightarrow 0} \frac{(\frac{1}{2} + h)^2 - (\frac{1}{2})^2}{h} = \lim_{h \rightarrow 0} \frac{h + h^2}{h} = \lim_{h \rightarrow 0} (1 + h) = 1.$$

With equal ease we can find the slope of the tangent at any point of the curve (x_1, y_1) , where $y_1 = x_1^2$. It is

$$\lim_{h \rightarrow 0} \frac{(x_1 + h)^2 - x_1^2}{h} = \lim_{h \rightarrow 0} \frac{2x_1h + h^2}{h} = \lim_{h \rightarrow 0} (2x_1 + h) = 2x_1.$$

Generalizing, let us try to express the slope of the tangent to any curve $y = f(x)$, at (x_1, y_1) . We imagine a chord through (x_1, y_1) and a neighboring point of the curve $(x_1 + h, y_1 + k)$. Since both points are on the curve, $y_1 = f(x_1)$ and $y_1 + k = f(x_1 + h)$ (Fig. 13). The slope of the chord is

$$\frac{k}{h} = \frac{f(x_1 + h) - f(x_1)}{h}.$$

We now define the slope of the tangent at (x_1, y_1) to be

$$\lim_{h \rightarrow 0} \frac{k}{h} = \lim_{h \rightarrow 0} \frac{f(x_1 + h) - f(x_1)}{h},$$

if this limit exists. There is still the problem of calculating the limit just

expressed. Chapter 3 will be devoted to this problem, particularly in connection with algebraic functions.

The limit,

$$m = \lim_{h \rightarrow 0} \frac{f(x_1 + h) - f(x_1)}{h},$$

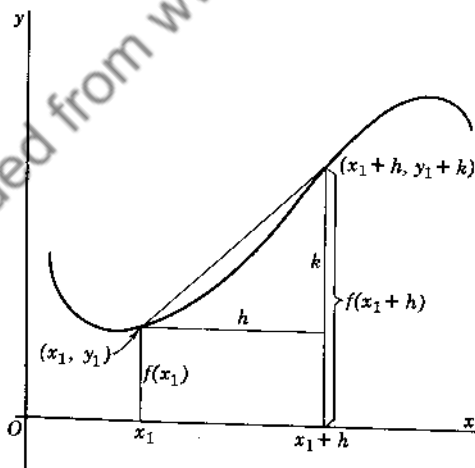


FIG. 13

defines the slope of the tangent to $y = f(x)$ at $x = x_1$. The tangent line at this point is the line

$$y - y_1 = m(x - x_1),$$

found from the point-slope form of the straight-line equation in analytic geometry. We have thus defined the slope of the tangent and the tangent line in such a way as to meet intuitive conceptions of a tangent.

EXAMPLE. Find the tangent line to the curve $y = 1/(x - 1)$ at $(2, 1)$.

SOLUTION. From the above definition, the slope is

$$\begin{aligned}\lim_{h \rightarrow 0} \left(\frac{1}{2 + h - 1} - \frac{1}{2 - 1} \right) / h &= \lim_{h \rightarrow 0} \frac{-h}{h(1 + h)} \\ &= \lim_{h \rightarrow 0} \frac{-1}{1 + h} = -1.\end{aligned}$$

The equation of the tangent is

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

EXERCISES

Find the slopes of the tangents to the curves in Exercises 1-10 at the points specified, and write the equations of the tangents. In each case form the quotient $\frac{f(x_1 + h) - f(x_1)}{h}$ and evaluate the limit of this quotient as $h \rightarrow 0$.

1. $y = x^2/4$, at $(2, 1)$; $(0, 0)$; $(-1, \frac{1}{4})$.
2. $y = -x^2$, at $(1, -1)$; $(-2, -4)$.
3. $y = 2x^2$, at $(1, 2)$; $(-1, 2)$.
4. $y = x^2/4 + 2x - 5$, at $(2, 0)$; $(0, -5)$.
5. $y = 4 - 3x - x^2$, at $(0, 4)$; $(1, 0)$.
6. $y = 1/x$, at $(1, 1)$; $(2, \frac{1}{2})$; $(\frac{1}{2}, 2)$.
7. $y = 1/x^2$, at $(1, 1)$; $(2, \frac{1}{4})$.
8. $y = \sqrt{x}$, at $(1, 1)$; $(4, 2)$.
9. $y = -\sqrt{x}$, at $(1, -1)$; $(9, -3)$.
10. $y = 1/\sqrt{x}$, at $(1, 1)$; $(4, \frac{1}{2})$.

11. Form the quotient $\frac{f(x_1 + h) - f(x_1)}{h}$ for $f(x) = \sin x$, $x_1 = 0$, and show that the limit to be calculated is $\lim_{h \rightarrow 0} \frac{\sin h}{h}$. It will be proved later that this limit is 1. Taking this for granted, find the slope of the tangent to the curve $y = \sin x$ at $(\pi, 0)$.

12. Form the quotient $\frac{f(x_1 + h) - f(x_1)}{h}$ for $f(x) = 2^x$, $x_1 = 0$. The limit which gives the slope of the tangent to the curve $y = 2^x$ at $(0, 1)$ will be later shown to be approximately .693. Taking this result for granted, show how to find the slope at any point (x_1, y_1) on the curve. Then find the slope for the points $(1, 2)$, $(2, 4)$, and $(-1, 1/2)$.

3

Derivatives

1. Derivatives. Following the lead suggested by our intuitive conception of a tangent, we arrived at a definition of the slope of a tangent, in Chapter 2, by means of $\lim_{h \rightarrow 0} \frac{f(x_1 + h) - f(x_1)}{h}$. In Chapter 1 we saw that any rate problem could be identified with the problem of the slope of a tangent, which may be considered as the graphical interpretation of a rate problem. Thus, if $v = f(u)$, we say that $\lim_{h \rightarrow 0} \frac{f(u_1 + h) - f(u_1)}{h}$ is the precise rate of change of v with respect to u , at $u = u_1$. This is a numerical definition that is independent of the particular interpretation of the limit as the slope of a tangent. We call this numerical limit the *derivative* of the function from which it is obtained, in accordance with the following definition.

DEFINITION. The derivative of the function $f(x)$ with respect to x , at the point $x = x_1$, is the number $\lim_{h \rightarrow 0} \frac{f(x_1 + h) - f(x_1)}{h}$, provided that this limit exists.

If the limit designated in this definition does not exist, the function $f(x)$ has no derivative at $x = x_1$. In this case it is impossible to specify a precise rate of change of the function at the given point. Clearly the derivative does not exist unless $f(x_1 + h) - f(x_1)$ tends to 0 as h tends to 0, for otherwise the function of h , $\frac{f(x_1 + h) - f(x_1)}{h}$, would increase beyond all bounds. That is, the derivative of a function does not exist unless the function is continuous at the point in question.

The phrase "with respect to x " draws attention to the fact that the independent variable, which can be varied at will, is x . Problems will arise in which there may be a question as to what variable is independent. For example, x may be a function of another variable and therefore vary only as a result of changes in the latter instead of independently. The need for care in specifying the variable with respect to which the derivative is found will be appreciated when other variables appear.

We emphasize that the derivative of $f(x)$ with respect to x , at $x = x_1$, is a definite fixed number. It is not a variable, and it is not an approximation inexactly conceived. Indeed, it is the precise number upon which a sequence of approximations, as described in Chapter 2, converges.

Derivatives were found in Chapter 2. Each of Exercises 1-10, Art. 13, Chapter 2, required the calculation of a derivative. There we found a derivative by the formation of the quotient $\frac{f(x_1 + h) - f(x_1)}{h}$ and then

the determination of the limit of this function of h , as required by the definition. Much of the labor involved in this process can be removed by the development of formulas by which derivatives of many functions can be written at sight. These formulas will be derived by applying the process called for in the definition to various kinds of functions.

The process of finding a derivative is called *differentiation*. To *differentiate* a function is to find a derivative of it. That part of calculus that centers around derivatives is called *differential calculus*. The basic method of differentiation is, of course, the direct application of the limit process required by the definition of a derivative, and we must be prepared to use this method to differentiate any function that does not fall within the scope of known formulas for derivatives, such as those about to be derived.

2. Derived Functions. Thus far in the process of finding derivatives we have first fixed x at some value, say x_1 . It was immaterial just what value of x was taken as x_1 , as long as it was on the range of definition of the function. This fact leads us to generalize the process in the following manner. For the function $f(x) = x^2$, for example, we write the derivative of $f(x)$ at x_1 as

$$\lim_{h \rightarrow 0} \frac{(x_1 + h)^2 - x_1^2}{h} = \lim_{h \rightarrow 0} (2x_1 + h) = 2x_1.$$

Since we could have let x_1 be any value of x whatever, we may drop the subscript and write the derivative of x^2 at any value of x ,

$$\lim_{h \rightarrow 0} \frac{(x + h)^2 - x^2}{h} = 2x.$$

We thus obtain a formula for the derivative of x^2 for any value of x . Thus, for $x = 1$, the derivative is $2 \cdot 1 = 2$; for $x = 2$, the derivative is $2 \cdot 2 = 4$; for $x = -2$, the derivative is $2(-2) = -4$; etc.

In a similar way, we can derive formulas for the derivatives of many of the functions that occur most frequently in practice. Each of these formulas expresses the derivative as a function of x . In common practice, the word "derivative" is used in two senses: (1) as a single number defined

by $\lim_{h \rightarrow 0} \frac{f(x_1 + h) - f(x_1)}{h}$, and (2) as the function of x from which the "derivative at a point $x = x_1$ " can be found by substitution of $x = x_1$. When there is danger of confusion it is best to use the term "derivative at $x = x_1$ " when the reference is to a single number, and "derived function" when the formula for the derivative is intended. In any event, it should be borne in mind that the fundamental idea of the derivative is based on the process of finding the limit of a function of h , namely $\frac{f(x + h) - f(x)}{h}$, while x is held fixed.

In order to reduce the work of differentiation to a minimum, it is important to learn the derived functions for many simple functions. We shall find, for example, that the knowledge of the derived function of x^n , where c is a constant and n is a positive integer, together with a rule for differentiating the sum of functions, will make it possible to differentiate any polynomial very easily. Without great difficulty we shall be able to derive further formulas by which we can differentiate all algebraic functions.

3. Notation. The expression $\lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$, which embodies the whole definition of a derivative, is also a symbol for a derivative. For convenience and conciseness, however, we need a more compact symbol. Several symbols for this purpose are in use, and it is well to be familiar with them.

When the given function is represented by some letter, as in an equation of the form $y = f(x)$, the derived function may be denoted by Dy , y' , or dy/dx .^{*} When we wish to express the derived function of $f(x)$ without interposing a letter to represent the dependent variable, it can be denoted by $Df(x)$, $f'(x)$, or $\frac{d}{dx}f(x)$.^{*} We shall make use of all these notations, but for a time we shall confine our usage to the first one, viz., Dy or $Df(x)$. This is read "dee y ," or "dee f of x ." It should not be mistaken for a product. In this symbol we may, if we wish, replace $f(x)$ by the explicit expression that defines it. Thus, if $f(x) = x^2 - x$, we may write $Df(x) = D(x^2 - x)$. We shall also use the symbol $D_x y$, inserting the subscript to call attention to the fact that the independent variable is x .

^{*} These notations are due to Cauchy, Lagrange, and Leibniz, respectively. There are many advantages in the use of the Leibniz notation, viz., $\frac{d}{dx}$, which will appear at a later point. At the start, however, a student often gets a wrong impression from this notation, and it seems best to avoid its use until the fundamental ideas are clear.

If we wish to denote the derivative of a function at a specified value of x , we may add to the symbol for the derived function a statement indicating the value specified. Thus $Df(x)_{x=x_1}$ means the derivative of $f(x)$ at $x = x_1$. Similarly, we may write $(Dy)_{x=x_1}$. Sometimes the concise notation $Df(x_1)$ is used. This symbol suggests the first step of the operation by which the derivative at $x = x_1$ is determined, namely, that of fixing x at x_1 , and then proceeding with the limiting process. In practice, however, we usually find the derived function first and then substitute in this function in order to get the derivative at $x = x_1$. The symbol $Df(x)_{x=x_1}$ suggests this order.

Other notations for the derivative at $x = x_1$ are $f'(x)_{x=x_1}$, or $f'(x_1)$; $\frac{d}{dx}f(x)_{x=x_1}$; $\left(\frac{dy}{dx}\right)_{x=x_1}$; and $(y')_{x=x_1}$. Sometimes these symbols are further abbreviated to $Df(x)_{x_1}$, etc.

4. Some Important Formulas. Let us first find the derived functions for some simple powers of x . In each case we imagine x to be fixed, even though we dispense with the subscript to emphasize this fact. The result is general, since the particular value at which x is fixed is not specified; that is, the result is the derived function.

(1) Let $f(x) = x$. Then $f(x + h) = x + h$, and $f(x + h) - f(x) = x + h - x = h$. Hence

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

and

$$\lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1.$$

Therefore

$$Df(x) = D(x) = 1.$$

This result is otherwise evident when we consider the derivative as the slope of the graph of the equation $y = x$.

(2) Let $f(x) = x^2$. We have already found the derived function in this case (Chapter 1, Art. 8, and Chapter 2, Art. 7, and Art. 13). It is

$$D(x^2) = 2x.$$

(3) Let $f(x) = x^3$. Then

$$\begin{aligned} f(x + h) &= (x + h)^3 - x^3 \\ &= (x + h - x)(x + h^2 + x \cdot \overline{x + h} + x^2) \\ &= h(x + h^2 + x \cdot \overline{x + h} + x^2) \end{aligned}$$

Hence

$$\frac{f(x+h) - f(x)}{h} = \frac{h}{h} (\overline{x+h^2} + x \cdot \overline{x+h} + x^2)$$

and

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} (\overline{x+h^2} + x \cdot \overline{x+h} + x^2).$$

As $h \rightarrow 0$, each of the three terms in the parentheses tends to x^2 , and therefore $D(x^3) = 3x^2$.

(4) Let $f(x) = x^n$. Then

$$\begin{aligned} f(x+h) - f(x) &= (x+h)^n - x^n \\ &= (x+h-x)(\overline{x+h^{n-1}} + \overline{x+h^{n-2}}x \\ &\quad + \overline{x+h^{n-3}}x^2 + \cdots + x^{n-1}),^* \end{aligned}$$

$$\frac{f(x+h) - f(x)}{h} = \frac{h}{h} (\overline{x+h^{n-1}} + \overline{x+h^{n-2}}x + \overline{x+h^{n-3}}x^2 + \cdots + x^{n-1}),$$

and

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} (\overline{x+h^{n-1}} + \overline{x+h^{n-2}}x + \overline{x+h^{n-3}}x^2 + \cdots + x^{n-1}).$$

As h tends to 0, each of the n terms in the parentheses tends to x^{n-1} , and therefore

$$D(x^n) = nx^{n-1}. \quad (D-1)^\dagger$$

Although we have proved this formula only for positive integral values of n , the formula nevertheless holds for all integral and fractional values of n , as will be shown later, except when $x = 0$ and $n < 1$. For these exceptions the formula has no meaning since x^{n-1} takes the form $1/0$. This formula is used very frequently, either by itself or in combination with other formulas, throughout calculus, and should be thoroughly familiar.[†]

Stated in words, formula D-1 is the following rule.

RULE. The derivative of the n th power of x is the $(n-1)$ st power of x multiplied by n .

This formula is often referred to as the *power formula* for differentiation.

* Since $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + b^{n-1})$.

† General formulas for differentiation of specific types of functions will be given numbers with the letter D prefixed, to which reference will be made throughout the book.

(5) Let $f(x) = c$. Then $f(x + h) - f(x) = c - c = 0$. Hence

$$\frac{f(x + h) - f(x)}{h} = 0,$$

$$\lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = 0,$$

and

$$D(c) = 0.$$

This fact is immediately evident when one considers the derivative as a rate of change of a function. Since the function is constant, its rate of change is 0. Also, the slope of the graph of $y = c$ is 0.

The result $D(1) = 0$ can be obtained as a special case from the power formula. Here we have $n = 0$, since $x^0 = 1$. Formula D-1 then gives $D(1) = D(x^0) = 0 \cdot x^{-1} = 0$, if $x \neq 0$.

5. Derivative of a Sum. The simplest and, at the same time, most important type of algebraic function is a polynomial, which is the algebraic sum of terms like cx^n , where c is any constant and n is a positive integer. Let us see how to differentiate a sum of two powers of x , e.g., $f(x) = x^3 + x^2$. We have

$$\begin{aligned} f(x + h) - f(x) &= (x + h)^3 + (x + h)^2 - x^3 - x^2 \\ &= (x^3 + h^3 - x^3) + (x^2 + h^2 - x^2) \end{aligned}$$

It is now easy to see that the remaining steps can be taken directly from Examples 2 and 3 of Art. 4, and that

$$D(x^3 + x^2) = D(x^3) + D(x^2) = 3x^2 + 2x. \quad (5.1)$$

The extension of this result to more terms offers no difficulty.

The generalization of the result in (5.1) to the sum of any functions that can be differentiated leads to an important theorem. Suppose, for example, that $f(x) = u(x) + v(x)$, where $u(x)$ and $v(x)$ are any functions whose derivatives can be found. Then

$$\begin{aligned} \frac{f(x + h) - f(x)}{h} &= \frac{u(x + h) + v(x + h) - u(x) - v(x)}{h} \\ &= \frac{u(x + h) - u(x)}{h} + \frac{v(x + h) - v(x)}{h}, \end{aligned}$$

and it follows that

$$Df(x) = D[u(x) + v(x)] = Du(x) + Dv(x). \quad (5.2)$$

The extension to any number of functions is immediate, and hence we have the following theorem.

THEOREM 1. The derivative of the algebraic sum of two or more functions is the algebraic sum of their derivatives.

A second theorem, which is equally fundamental, enables us to differentiate the product of a constant by any function. It is almost obvious and is proved in much the same way as Theorem 1. Thus, if $f(x) = cu(x)$,

$$\frac{f(x+h) - f(x)}{h} = c \frac{u(x+h) - u(x)}{h},$$

and it follows that

$$Df(x) = D[cu(x)] = cDu(x). \quad (5.3)$$

Thus, we have Theorem 2.

THEOREM 2. The derivative of the product of a constant by a function is the product of the constant multiplied by the derivative of the function.

With these two theorems and the power formula, we can differentiate any polynomial at sight.

EXAMPLE 1.

$$\begin{aligned} D(2x^2 + 3x - 1) &= D(2x^2) + D(3x) - D(1) \\ &= 2D(x^2) + 3D(x) - D(1) \\ &= 2 \cdot 2x + 3 \cdot 1 - 0 \\ &= 4x + 3. \end{aligned}$$

EXAMPLE 2.

$$\begin{aligned} D(4x^4 - 5x^3 - 3x^2 + 7x + 5) \\ &= D(4x^4) - D(5x^3) - D(3x^2) + D(7x) + D(5) \\ &= 4 \cdot 4x^3 - 5 \cdot 3x^2 - 3 \cdot 2x + 7 \cdot 1 + 0 \\ &= 16x^3 - 15x^2 - 6x + 7. \end{aligned}$$

EXERCISES

Find the derived functions of the functions in Exercises 1–20.

1. x^3 .	2. x^8 .	3. x^{15} .	4. 200.	5. $4x^2$.
6. $2x^7$.	7. $-5x^3$.	8. $-x^4$.	9. $x - 3$.	10. $x^2 + \frac{1}{3}$.
11. $8x^2 - 3$.			12. $3x^3 - 4x^2 + x + 2$.	
13. $x^3/3 + x^2/2 + 1$.			14. $\sqrt{2}x^2 - x/\sqrt{2} - 1$.	
15. $ax^2 + bx + c$ (a , b , and c = constants).			16. $3x^5 - 5x^3$.	
17. $(x-1)(x+1)$.			18. $(x-1)(x+3)$.	
19. $2 - 4x - 6x^2 + 5x^3$.			20. $5 + x - x^2/2 + x^3/6$.	

Find the derivatives of each of the functions in Exercises 21–25 at the points specified, by first finding the derived function and then inserting the given values.

21. $f(x) = 3x^3 - 10x^2 + 4$, at $x = 0$; $x = 1$; $x = -1$.
 22. $f(x) = 2x - x^3$, at $x = -1$; $x = 2$.

23. $f(t) = 3 + 4t - 6t^2 + 7t^3$, at $t = 0$; $t = 1$; $t = 5$.

24. $f(u) = u^4 - 2u^2 + 1$, at $u = 1$; $u = -1$; $u = 0$.

25. $f(r) = r^5/3 + r^4/2 - 4r$, at $r = 1$; $r = -1$; $r = 2$.

Find the slope of the tangent to each of the curves in Exercises 26-30 at the abscissa specified, and write the equation of the tangent.

26. $y = 4 - x^2$, at $x = 2$.

27. $y = x^3/6$, at $x = -1$.

28. $y = 4x - x^3$, at $x = 1$.

29. $y = x^2 - 6x + 2$, at $x = 3$.

30. $y = x^3 - 3x + 2$, at $x = 1$, and at $x = -1$.

31. It is found in physics that the number of feet s that a body falls in t seconds, from a position of rest, is $s = 16t^2$. Here the derivative of s with respect to t , at any value of t , is the rate of change of s per unit time at this instant, i.e., the instantaneous velocity (see Chapter I, Art. 7). Find the velocity at the end of 1 sec; 2 sec; 3 sec; 10 sec.

32. If a body is thrown upward from the earth with an initial velocity of 100 ft per sec, its distance from the earth in feet is given by the equation $s = 100t - 16t^2$, where t is the time elapsed, measured in seconds. Find the velocity when $t = 1$; $t = 2$, $t = 3$, $t = 4$; $t = 6$. Explain the significance of the sign of the result.

33. If a body moves so that its distance, s feet, from the starting point at the end of t seconds is given by the formula $s = 80t - 16t^2$, find its velocity at the end of 1 sec; 2 sec; 4 sec.

34. The acceleration of a moving body is the time rate of change of velocity. If the velocity is related to the time by the equation $v = 2t^2$, find the acceleration when $t = 1$; $t = 2$.

35. If $s = f(t)$ is the equation of motion of a body moving in a straight line, the derived function $Df(t)$ gives the velocity at any time t . We may write $v = Df(t) = \phi(t)$. Since the rate of change of velocity is the acceleration a , we have $a = Dv = D\phi(t)$. Show that the acceleration of the body in Exercise 31 is constant, and that the same is true in Exercise 32 and in Exercise 33.

36. Use Exercise 35 to find the acceleration of a body that moves in accordance with the equation $s = 3t^3 - 8t$, when $t = 1$; when $t = 2$.

37. The consumption of coal by a locomotive depends on the speed maintained, in accordance with the formula $T = 0.005v^2$, in which v is the speed in miles per hour and T is the number of tons of coal consumed per hour. Find the rate at which the consumption of coal increases per unit increase in speed, when the speed is 40 mph.

38. The maximum range of a projectile shot from a certain gun is related to the muzzle velocity by the formula $R = 0.024v^2$. Find the rate of change in the maximum range, per unit change in the muzzle velocity, when $v = 1500$.

39. Differentiate each of the following by direct application of the definition of the derivative, and show that the power formula holds in each case:

(a) $f(x) = x^{-1} = 1/x$.

(b) $f(x) = x^{-2} = 1/x^2$.

(c) $f(x) = x^{\frac{1}{2}} = \sqrt{x}$.

(d) $f(x) = x^{\frac{1}{3}} = \sqrt[3]{x}$.

40. Differentiate each of the following by direct application of the definition of the derivative:

(a) $f(x) = 1/(x - 1)$.

(b) $f(t) = (t + 1)/(t - 1)$.

(c) $f(s) = s/(1 - s^2)$.

(d) $f(x) = 1/\sqrt{x}$.

6. Derivatives of Products and Quotients. Theorem 1 of Art. 5, governing differentiation of a sum, was almost self-evident. A rule for differentiating a product of two functions is less obvious. Let

$$f(x) = u(x) \cdot v(x),$$

where both $u(x)$ and $v(x)$ are functions that we can differentiate. Then

$$f(x+h) - f(x) = u(x+h) \cdot v(x+h) - u(x) \cdot v(x).$$

Subtracting and adding $u(x) \cdot v(x+h)$, we have

$$\begin{aligned} f(x+h) - f(x) &= u(x+h) \cdot v(x+h) - u(x) \cdot v(x+h) + u(x) \cdot v(x+h) - u(x) \cdot v(x) \\ &= v(x+h)[u(x+h) - u(x)] + u(x)[v(x+h) - v(x)]. \end{aligned}$$

Then

$$\frac{f(x+h) - f(x)}{h} = v(x+h) \frac{u(x+h) - u(x)}{h} + u(x) \frac{v(x+h) - v(x)}{h}.$$

Applying the theorems on limits of Art. 9, Chapter 2,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} &= \left[\lim_{h \rightarrow 0} v(x+h) \right] \cdot \left[\lim_{h \rightarrow 0} \frac{u(x+h) - u(x)}{h} \right] \\ &\quad + u(x) \cdot \left[\lim_{h \rightarrow 0} \frac{v(x+h) - v(x)}{h} \right]. \end{aligned}$$

Hence

$$Df(x) = D[u(x) \cdot v(x)] = v(x) \cdot Du(x) + u(x) \cdot Dv(x). \quad (6.1)$$

Equation 6.1 is called the *product formula* and is expressed verbally by the theorem.

THEOREM I. The derivative of the product of two functions is the sum of the products of each of the functions by the derivative of the other.

Formula 6.1 can be easily extended to products of three or more factors. Thus

$$\begin{aligned} D(uvw) &= uv Dw + w D(uv) = uv Dw + w(uDv + vDu) \\ &= uv Dw + wu Dv + vw Du. \end{aligned}$$

In order to differentiate algebraic fractions we need a formula for the derivative of a quotient of two functions. Let $f(x) = u(x)/v(x)$. Then

$$\begin{aligned} f(x+h) - f(x) &= \frac{u(x+h)}{v(x+h)} - \frac{u(x)}{v(x)} \\ &= \frac{u(x+h) \cdot v(x) - v(x+h) \cdot u(x)}{v(x) \cdot v(x+h)}. \end{aligned}$$

Subtracting and adding $u(x) \cdot v(x)$ in the numerator, this fraction becomes

$$\frac{u(x+h) \cdot v(x) - u(x) \cdot v(x) - v(x+h) \cdot u(x) + u(x) \cdot v(x)}{v(x) \cdot v(x+h)},$$

or

$$\frac{v(x)[u(x+h) - u(x)] - u(x)[v(x+h) - v(x)]}{v(x) \cdot v(x+h)}.$$

Dividing both members by h , with the division applied to the numerator on the right-hand side, we have

$$\frac{f(x+h) - f(x)}{h} = \frac{v(x) \cdot \frac{u(x+h) - u(x)}{h} - u(x) \cdot \frac{v(x+h) - v(x)}{h}}{v(x) \cdot v(x+h)}.$$

In order to find $\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$ we apply the theorems on limits, of Art. 9, Chapter 2, on the right-hand side, and obtain

$$Df(x) = D \frac{u(x)}{v(x)} = \frac{v(x) \cdot Du(x) - u(x) \cdot Dv(x)}{[v(x)]^2}. \quad (6.2)$$

This is the *quotient formula* for differentiation. Expressed verbally, it is Theorem 2.

THEOREM 2. The derivative of a fraction, i.e., the quotient of two functions, is equal to the denominator multiplied by the derivative of the numerator minus the numerator multiplied by the derivative of the denominator, all divided by the square of the denominator.

EXAMPLE 1.

$$\begin{aligned} D \frac{x^2 - 1}{x^2 + 1} &= \frac{(x^2 - 1)D(x^2 + 1) - (x^2 + 1)D(x^2 - 1)}{(x^2 + 1)^2} \\ &= \frac{(x^2 - 1)(2x) - (x^2 + 1) \cdot 2x}{(x^2 + 1)^2} = -2 \cdot \frac{x^2 + x + 1}{(x^2 + 1)^2}. \end{aligned}$$

As a special case of this formula we may let $u(x) = c$:

$$D \frac{c}{v(x)} = \frac{-c Dv(x)}{[v(x)]^2}. \quad (6.3)$$

EXAMPLE 2.

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

When the numerator is constant, formula 6.3 should be used in preference to (6.2).

We are now in a position to differentiate any rational function, i.e., any function that can be expressed as the quotient of two polynomials.

EXERCISES

Find the derived function of each of the functions in Exercises 1-10.

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

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

$$3. \frac{x^2+1}{x+1}.$$

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

$$5. 1/(x+1).$$

$$6. x+1/x.$$

$$7. x^2-1/x^2.$$

$$8. (x^2+x+1)(x^2-x+1).$$

$$9. (1+1/x)(x^2-1).$$

$$10. x^5(2x^3+1)(x^2-2).$$

11. Find $D(1/x)$, $D(1/x^2)$, and $D(1/x^n)$. Then show that the power formula (2-1) holds for negative integral powers.

12. Find the slope of the tangent to each of the following curves at the point specified: (a) $y = 1/x$, at $(\frac{1}{2}, 2)$. (b) $y = 1/x^2$, at $(-1, 1)$. (c) $y = (x+1)/(x^2-4)$, at $(1, -\frac{3}{2})$. (d) $y = x - 108/x^2$, at $(6, 3)$.

7. Related Rates. We have seen that, when y is a function of x , $y = f(x)$, the rate of change of y with respect to x , i.e., per unit change in x , is given by the derivative. A slightly different kind of problem is illustrated by the following example. Suppose a stone is dropped in a pool of still water. A set of concentric ripples is started, the disturbed area being an expanding circle. From an observation of the rate of increase in the radius how can we find the rate of increase of the disturbed area? Here A is a function of r ; viz., $A = \pi r^2$. We think of r as a function of t , and, hence, A is also a function of t . Let us suppose that r increases constantly at 3 ft per sec and, hence, $r = 3t$ and $Dr = 3$. We wish to find $D_t A$. The subscript t is added to the derivative notation to indicate that it is the derivative with respect to t that is wanted, i.e., the time rate of change of A . From the relation $A = \pi r^2$, we have $D_r A = 2\pi r$. That is, the area increases $2\pi r$ units per unit increase in r . Now, if r increased just 1 ft per sec, this would mean that the area must increase $2\pi r$ sq ft per sec. Since r increases 3 ft per sec, the area must increase 3 times as fast, or $2\pi r \cdot 3$ sq ft per sec. Thus, to find the number of units increase in the area per second, we multiply the number of units increase in the area per unit increase in r by the number of units increase in r per second. That is,

$$\begin{aligned} D_t A &= D_r A \cdot D_t r \\ &= 2\pi r \cdot D_t r = 2\pi r \cdot 3. \end{aligned} \quad (7.1)$$

At the moment when r reaches 12 ft, the rate of increase in A is $2\pi \cdot 12 \cdot 3$, or 72π sq ft per sec.

This solution is easily checked by substituting $r = 3t$ in $A = \pi r^2$, obtaining $A = 9\pi t^2$. Hence, $D_t A = 18\pi t$. At the end of 4 sec, $r = 12$ and $D_t A = 18\pi \cdot 4 = 72\pi$. This is, of course, an exceptionally simple example. The calculation of the rate by the substitution process is usually much less convenient than the use of the former method. We shall consider another example.

EXAMPLE. A conical funnel is 12 in. across the top and 12 in. deep. A liquid is flowing in at the rate of 6 cu in. per sec and flowing out at the rate of 3 cu in. per sec. At what rate is the surface of the liquid rising when the liquid is 8 in. deep?

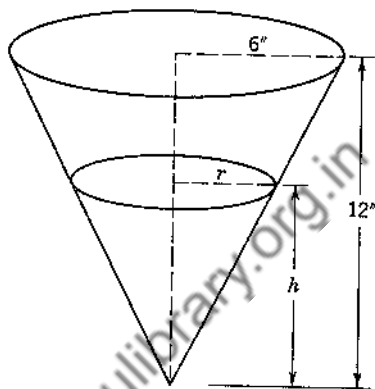


FIG. 1

SOLUTION. From the formula for the volume of a cone, the volume of the liquid at any moment is $V = \pi r^2 h / 3$, where h is the depth and r is the radius of the circular surface (Fig. 1). From similar triangles, $r/h = 6/12$, i.e., $r = \frac{1}{2}h$. Hence, $V = \pi h^3 / 12$. We now consider both V and h as functions of t , the number of seconds the liquid has been filling the funnel. Reasoning as in the illustration above, we conclude that $D_t V = D_h V \cdot D_t h$. That is,

$$D_t V = \frac{3\pi h^2}{12} \cdot D_t h. \quad (7.2)$$

Since the liquid flows in at 6 cu in. per sec, and flows out at 3 cu in. per sec, V increases at the rate of 3 cu in. per sec; i.e., $D_t V = 3$. We require $D_t h$ at the instant when $h = 8$. Therefore, we place $D_t V = 3$ and $h = 8$ in (7.2), obtaining

$$3 = \frac{3\pi}{12} \cdot 64 \cdot D_t h$$

whence

$$D_t h = \frac{3}{16\pi}.$$

That is, the surface is rising at the rate of $3/16\pi$, or about .06 in. per sec.

EXERCISES

1. Find the rate of change of the area of a square when the side is 8 in. and increasing at the rate of 2 in. per sec.
2. Find the rate of change of the volume of a cube when the edge is 8 in. and increasing at the rate of 2 in. per sec.

3. At what rate is the area of the surface of a sphere changing when the radius is 2 in. and increasing at the rate of $\frac{1}{2}$ in. per sec? [$A = 4\pi r^2$.]
4. At what rate is the volume of a sphere changing when the radius is 2 in. and increasing at the rate of $\frac{1}{2}$ in. per sec? [$V = \frac{4}{3}\pi r^3$.]
5. If the ordinate of a point moving along the parabola $y^2 = 12x$ increases uniformly at the rate of 2 in. per sec, at what rate is the abscissa changing when $y = 6$? when $y = 12$?
6. A man 6 ft tall walks away from a lamp 10 ft high at the rate of 4 mph. How fast does the shadow of his head move away from the base of the lamppost? How fast is the shadow lengthening?
7. Sand is being poured on the ground at the rate of 5 cu ft per min and forms a pile which has the shape of a right circular cone with altitude equal to the radius of the base. How fast is the height of the pile increasing when the height is 3 ft?
8. A perfect gas confined in a container and kept at a constant temperature obeys Boyle's law: $pv = a$ constant, where v represents the volume, and p represents the pressure which the gas exerts on every square inch of interior surface of the container. If, at a certain time, the volume is 40 cu in. while the pressure is 50 lb per sq in. and increasing at the rate of 2 lb per sq in. per sec, find how fast the volume is changing.
9. The side of an equilateral triangle is 24 in. long and increasing at the rate of 3 in. per hr. How fast is the area increasing?
10. A trough whose cross section is an equilateral triangle is 12 ft long and 18 in. across the top. If water is running into the trough at the rate of .5 cu ft per sec, at what rate is the surface rising when the trough is half full?

8. Compound Functions. The Chain Rule. The generalization of the method of related rates discussed in Art. 7 leads to a theorem of far-reaching importance. Let the function $u = \phi(x)$ possess a derivative $D_x u$ at all points in an interval $a \leq x \leq b$. As x varies in this interval, let u take on the values in an interval $a' \leq u \leq b'$, and let the function $y = f(u)$ possess a derivative $D_u y$ throughout this interval. It follows that the *compound function*, or function of a function, $y = f[\phi(x)]$ is a function of x in the interval $a \leq x \leq b$. We shall prove the following theorem.

THEOREM. If $y = f(u)$ and $u = \phi(x)$, subject to the above conditions, the function $y = f[\phi(x)] = F(x)$ possesses a derivative throughout the interval $a \leq x \leq b$, and

$$D_x y = D_u y \cdot D_x u. \quad (8.1)$$

Formula 7.1 is clearly a special case of formula 8.1, where x is identified with a measurement of time t . In order to prove this theorem we write

$$D_x y = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{f[\phi(x+h)] - f[\phi(x)]}{h},$$

provided that this limit exists. If we write $\phi(x) = u$ and $\phi(x+h) = u+k$, we have

$$\frac{F(x+h) - F(x)}{h} = \frac{f(u+k) - f(u)}{h}. \quad (8.2)$$

Here we take h different from 0, although we cannot guarantee that k will be different from 0. If k is different from 0, however, we can write

$$\lim_{k \rightarrow 0} \frac{f(u+k) - f(u)}{k} = D_u y.$$

We can be sure that this limit exists, since $f(u)$ is assumed to have a derivative throughout the interval. It follows that there is a function of the nonzero variable k , $\varepsilon(k)$, that tends to 0 as k tends to 0 and satisfies the equation

$$\frac{f(u+k) - f(u)}{k} = D_u y + \varepsilon(k) \quad (8.3)$$

(see Art. 8, Chapter 2). Hence,

$$f(u+k) - f(u) = [D_u y + \varepsilon(k)]k. \quad (8.4)$$

If we now extend the range of $\varepsilon(k)$ to include $k = 0$ by putting $\varepsilon(k) = 0$, equation 8.4 holds whether or not $k = 0$. Therefore, by substitution from (8.4) in equation 8.2,

$$\begin{aligned} \frac{F(x+h) - F(x)}{h} &= [D_u y + \varepsilon(k)] \frac{k}{h} \\ &= [D_u y + \varepsilon(k)] \frac{\phi(x+h) - \phi(x)}{h}, \end{aligned}$$

and

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} &= D_u y \cdot \left[\lim_{h \rightarrow 0} \frac{\phi(x+h) - \phi(x)}{h} \right] \\ &\quad + \left[\lim_{h \rightarrow 0} \varepsilon(k) \right] \cdot \left[\lim_{h \rightarrow 0} \frac{\phi(x+h) - \phi(x)}{h} \right]. \end{aligned}$$

Now, k , which is equal to $\phi(x+h) - \phi(x)$, tends to 0 as h tends to 0, since $\phi(x)$ is continuous, and hence $\lim_{h \rightarrow 0} \varepsilon(k) = \lim_{k \rightarrow 0} \varepsilon(k) = 0$. Therefore

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = D_u y \cdot \left[\lim_{h \rightarrow 0} \frac{\phi(x+h) - \phi(x)}{h} \right]$$

or

$$D_x y = D_u y \cdot D_x u. \quad (8.1)$$

Formula 8.1 is called the *chain rule* of differentiation, linking together the separate derivatives of two functions that are compounded, to form the derivative of the compound function. It will be called into use more and more as we extend differentiation to functions of greater complexity. This formula can be extended without difficulty to include more than one intermediate variable. Thus, for example, if $y = \phi(u)$, $u = \psi(v)$, and $v = f(x)$, then $y = \phi[\psi[f(x)]] = F(x)$, and $D_x y = D_u y \cdot D_v u \cdot D_x v$.

Formulas for derivatives are generalized by means of the chain rule. Thus, we broaden the power formula to deal with functions of the type $y = u^n$, where u is any function whose derivative is known. By (3.1), we have

$$D_x y = nu^{n-1} D_x u. \quad (3.1a)$$

For example, if $y = (2x^3 - 1)^5$, we think of $2x^3 - 1$ as u . Then

$$y = u^5, \quad D_x y = 5u^4, \quad \text{and} \quad D_x u = 6x^2,$$

so that

$$D_x y = 5u^4 \cdot 6x^2 = 5(2x^3 - 1)^4 \cdot 6x^2.$$

With experience, we find it unnecessary to write in the substitution of u for $2x^3 - 1$.

By use of the quotient formula we found that $D(1/x) = -1/x^2$, and $D(1/x^n) = -n/x^{n+1}$. We can now write $D_x(1/u) = (-1/u^2) \cdot D_x u$, and $D_x(1/u^n) = (-n/u^{n+1}) D_x u$.

EXAMPLE 1. Find $D_x \left[\frac{x+1}{x-1} \right]^2$.

SOLUTION. Writing $u = \frac{x+1}{x-1}$, we have $D_x u = \frac{-2}{(x-1)^2}$, by the quotient formula. Hence

$$D_x \left[\frac{x+1}{x-1} \right]^2 = D_x u^2 = 2u \cdot D_x u = 2 \left[\frac{x+1}{x-1} \right] \cdot \frac{-2}{(x-1)^2} = \frac{-4(x+1)}{(x-1)^3}.$$

EXAMPLE 2. Find $D[1/(x+3)^3]$.

SOLUTION. Writing $u = x+3$, we have $D_x u = 1$,

$$\begin{aligned} D[1/(x+3)^3] &= D_x(1/u^3) = (-3/u^4) \cdot D_x u \\ &= -3/(x+3)^4. \end{aligned}$$

EXERCISES

Find $D_x y$ in each of the following exercises.

1. $y = u^3$, $u = 6x^3 - 3x^2 + 4x + 5$.

2. $y = 1/u$, $u = 5x^2 + 2$.

3. $y = 2u^3$, $u = 1/x$.

4. $y = 2u^2 - 3u + 4$, $u = \frac{x^2 - 1}{x^2 + 1}$.

5. $y = 3/u$, $u = (3x^2 - 2x + 1)^3$.

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

7. $y = (2-x)^2/(3+x)^3$.

8. $y = (1 - 1/x)^4$.

9. $y = \left(\frac{x-1}{x+1} \right)^3$.

10. $y = 2x^2 - 1 - 1/(2x^2 - 1)$.

9. Inverse Forms of a Functional Relation. The study of many functional relations is facilitated by approaching them through their *inverses*. The student is acquainted with numerous pairs of inverse relations. For example, $y = x^2$ and $x = \sqrt{y}$ is such a pair; so is $y = \log_{10} x$ and $x = 10^y$; and $y = \sin x$ and $x = \sin^{-1} y$ (or arc $\sin y$). In each of these examples, within suitably restricted ranges of the variables, the same correspondence between x and y is expressed by two forms, one of which is called the *inverse* of the other. On a graph, the correspondence between the variables can be read in two different orders. Thus, Fig. 2 shows the graph of $y = x^2$ for $x > 0$. This is also the graph of $x = \sqrt{y}$ for $y > 0$.

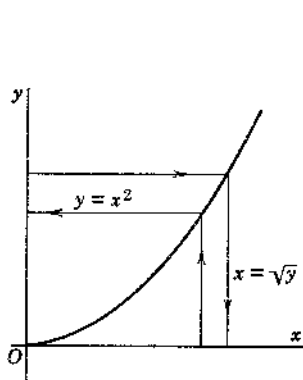


FIG. 2

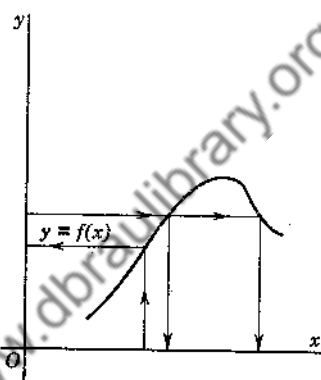


FIG. 3

We may start with a given value of x and read the corresponding value of y , or, in the inverse order, start with a value of y and read the corresponding value of x .

Consideration of the graphical representation shows that both of two inverse relations are single-valued only if the graph continually rises or continually falls. Thus, in Fig. 3, whereas a value of x gives rise to a single value of y , a given value of y yields two different values of x . If, on the other hand, the curve continually rises as x increases, representing a *monotonically increasing* function, $f(x)$; or, if it continually falls as x increases, representing a *monotonically decreasing* function; and, if $f(x)$ is continuous, there is a single-valued inverse relation $x = g(y)$.

The distinction between a pair of inverse relations amounts to a difference in point of view as to which of the two variables is independent. This difference may be indicated symbolically by writing $y = f(x)$ and $x = g(y)$. Of course, if, in either of these equations, we substitute for the independent variable the expression to which this variable is set equal in the other equation, the result is a simple identity in one

variable. Thus, if $y = x^2$, then $x = \sqrt{y}$, and if $\sqrt{y}$ replaces x in the first equation,

$$y = (\sqrt{y})^2 = y.$$

Similarly, if $y = \log_{10} x$, then $x = 10^y$, and by substitution, $y = \log_{10}(10^y) = y$. In general, if $y = f(x)$ and $x = g(y)$ are inverse relations, then

$$f[g(y)] = y, \quad \text{and} \quad g[f(x)] = x.$$

For a monotonically increasing function or a monotonically decreasing function we have the choice of two inverse representations. Often, one of these forms is more readily dealt with than the other. Thus, powers are more easily handled than roots, and, hence, $x = y^2$ seems a more approachable relation than $y = \sqrt{x}$. In this example, and in many others, the derivative of one form is more easily found than that of its inverse. We shall find a simple relation that always connects the derivative of a function with that of its inverse. This can be done with the help of the chain rule. For example, in order to obtain $D_x y$ from the relation $y = \sqrt{x}$, we can write the compound function, $y = \sqrt{x}$, $x = y^2$, which amounts to the identity $y = y$. In this compound relation, $D_y y = D_x y \cdot D_y x$. But, since $D_y y = 1$, and, from the equation $x = y^2$, $D_y x = 2y$, it follows that $1 = 2y \cdot D_x y$. Therefore, $D_x y = 1/(2y) = 1/(2\sqrt{x})$.

In general, let $y = f(x)$ and $x = g(y)$. Applying the chain rule, $D_y y = D_x y \cdot D_y x$. Since $D_y y = 1$, we have $1 = D_x y \cdot D_y x$, or

$$D_x y = 1/D_y x. \quad (9.1)$$

Expressed verbally, equation 9.1 is the following theorem.

THEOREM. The derivative of a function is equal to the reciprocal of the derivative of its inverse.

EXAMPLE 1. Find the slope of the graph of $y = \sqrt{x}$ at $x = \frac{1}{16}$, $y = \frac{1}{4}$.

SOLUTION. The inverse form of this relation is $x = y^2$. Making use of (9.1), we have $D_x y = 1/D_y x = 1/(2y)$. At $y = \frac{1}{4}$, $D_x y = 1/(2 \cdot \frac{1}{4}) = 2$, and the slope at the point in question is 2.

EXAMPLE 2. Find the derived function if $y = \sqrt[5]{x}$.

SOLUTION. The inverse form of this relation is $x = y^5$. Hence, making use of (9.1),

$$D_x y = 1/D_y x = 1/(5y^4) = 1/(5\sqrt[5]{x^4}).$$

Since this result is equal to $\frac{1}{5}x^{-\frac{4}{5}}$, we see that $D(x^{\frac{1}{5}}) = \frac{1}{5}x^{\frac{1}{5}-1}$, in agreement with the power formula.

10. Inverse Functions. Since it is the custom to represent the independent variable on the horizontal axis of rectangular coordinates, let us

try to discover the relationship between the graphs of inverse relations when this is done. For this purpose we interchange the x - and y -axes (Fig. 4). If we now let the independent variable and the dependent variable be x and y , respectively, for both relations, we see that the graphs

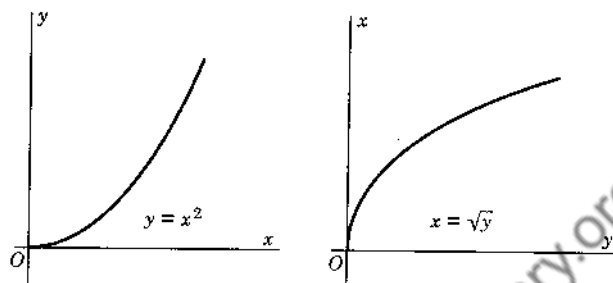


FIG. 4

of inverse relations are simply reflections of one another in the line $x = y$ (Fig. 5). If we place the emphasis on the forms of the relationships, disregarding the names of the variables, we still consider these relations as inverses. That is, if $f[g(y)] = y$, then we consider $y = f(x)$ and $y = g(x)$ as inverse relations, or inverse functions, although they no longer represent identically the same correspondence between x and y , as in Art. 9. Thus, $y = x^2$ and $y = \sqrt{x}$, $y = \log_{10} x$ and $y = 10^x$, and $y = \sin x$ and $y = \sin^{-1} x$ are considered as pairs of inverse functions.

11. The Power Formula for Rational Exponents.

Formula $D-1$ was derived for positive integral exponents in Art. 4. It was extended to negative integral exponents, by the help of the quotient formula, in Exercise 11, Art. 6. We now extend it to all rational exponents by means of the chain rule and the rule for differentiating inverse functions.

Let $y = x^{p/q}$, where p and q are integers. This can be expressed by the compound relation,

$$y = u^p, u = x^{1/q}.$$

Applying the chain rule,

$$D_x y = p u^{p-1} D_x u.$$

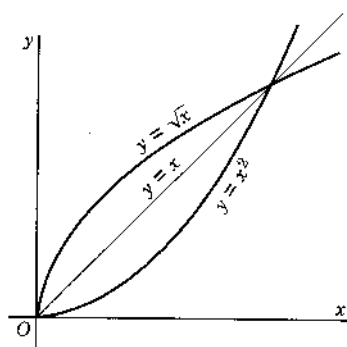


FIG. 5

Now, $D_x u$ can be found from $x = u^q$, which is the inverse of $u = x^{1/q}$. Thus, by (9.1),

$$D_x u = 1/D_u x = 1/(qu^{q-1}).$$

Hence,

$$D_x y = pu^{p-1}/(qu^{q-1}) = \frac{p}{q} u^{p-q}.$$

Since $u = x^{1/q}$, this becomes

$$Dy = \frac{p}{q} x^{\frac{1}{q}(p-q)} = \frac{p}{q} x^{\frac{p}{q}-1}.$$

This result agrees with formula D-1, with n replaced by p/q .

EXERCISES

1. Find $D_x y$ from each of the following:

(a) $x = y^2$.

(b) $x = 2y^2 - 3y + 2$.

(c) $x = 5y^3 + 2y^2$.

(d) $x = y + 1/y$.

(e) $x = 4/y^3$.

(f) $x = (y+1)/(1-y)$.

2. Find the slope of each of the following curves at the point specified:

(a) $x = \frac{1}{2}y^2$, at $(1, 2)$.

(b) $x = y^4 - 6y$, at $(-5, 1)$.

(c) $xy^2 = 12$, at $(3, 2)$.

(d) $y = x^{\frac{2}{3}}$, at $(1, 1)$.

(e) $y = x^{\frac{3}{2}}$, at $(4, 8)$.

(f) $2y^2 = x^3$, at $(2, 2)$.

3. Find the angles of intersection of the parabolas $y^2 = 6x$ and $x^2 = 6y$. [Hint:

Find the slope of each at the intersection points, and use the formula $\tan \alpha = \frac{m_1 - m_2}{1 + m_1 m_2}$.]

4. Find the angle of intersection of the cubical parabola $y^3 = x$ and the hyperbola $xy = 6$.

5. Use the general power formula D-1a, with the result of Art. 11, to find $D_x y$ from each of the following:

(a) $y = \sqrt{x^2 - 1}$.

(b) $y = (x^2 - 1)^{\frac{3}{2}}$.

(c) $y = 1/\sqrt{x^2 - 1}$.

(d) $y = 1/(x^2 - 1)^{\frac{3}{2}}$.

(e) $y = \sqrt{4 - x^2}$.

(f) $y = 1/\sqrt{a^2 - x^2}$.

12. Functions Expressed Parametrically. A compound functional relation employs a third variable as an intermediate variable. An auxiliary variable often appears in a slightly different way, with each of two variables expressed in terms of a third. In physical problems the third variable often represents time, although it can stand for any other quantity. Thus, if a body moves along a curve in a plane, each of its coordinates is a function of time: $y = f(t)$, $x = g(t)$. The auxiliary variable used in this way is called a *parameter*, and a relation between x and y is implied by the parametric equations that connect x and y with the parameter.

The representation of a functional relation by parametric equations can be considered a special case of a compound function. For, if $y = f(t)$ and $x = g(t)$, then $y = f(t)$, $t = \phi(x)$, in which $t = \phi(x)$ is the inverse relation to $x = g(t)$. By use of the chain rule and the rule for differentiating an inverse function, we then have

$$D_x y = D_t y \cdot D_x t = D_t y / D_t x. \quad (12.1)$$

EXAMPLE. If $y = u^2$, and $x = (u - 1)/(u + 1)$, find $D_x y$.

SOLUTION. Differentiating these functions, we have $D_u y = 2u$, $D_u x = 2/(u + 1)^2$. Hence, by (12.1),

$$D_x y = \frac{2u}{2/(u + 1)^2} = u(u + 1)^2.$$

EXERCISES

Find $D_x y$ in each of Exercises 1-5.

1. $y = u^2$, $x = u^3$.

2. $y = \sqrt{u}$, $x = 1/u$.

3. $y = t^2 + t$, $x = t^3 - t^2 + 2$.

4. $y = \sqrt{1 - t^2}$, $x = \sqrt{1 + t^2}$.

5. $y = (u^2 - 1)/(u^2 + 1)$, $x = 2u/(u^2 + 1)$.

6. Find the equation of the tangent to the curve $y = t^2$, $x = 3t - 1$, at the point (2, 1).

7. Find the equation of the tangent to the curve $y = 1/(t + 1)$, $x = t^2$, at the point where $t = -\frac{3}{2}$.

8. If an airplane flying horizontally at the rate of 150 mph drops a bomb, the path of the bomb has the parametric equations $y = -16t^2$, $x = 220t$, where the x -axis coincides with the path of the airplane, x and y are measured in feet, and t represents the number of seconds after the bomb is released. (These equations would be strictly accurate only in a vacuum.) Find the direction in which the bomb is moving 2 sec after it is dropped; 10 sec after it is dropped. Find the direction at which the bomb strikes the earth if the airplane was 14,400 ft high when the bomb was released. How fast is the bomb moving downward when it strikes?

9. A ball thrown with a speed of 50 ft per sec when it leaves the hand, in a direction inclined at 30° to the horizontal, moves so that its height t seconds after it is thrown is $y = 25t - 16t^2$. At the same time, its horizontal distance from the thrower is $x = 25\sqrt{3}t$. Find the direction of its motion at the end of $\frac{1}{2}$ sec; at the end of $\frac{2.5}{32}$ sec; when it strikes the ground.

13. Functions Expressed Implicitly. It is often convenient to represent a functional relation by an *implicit* equation; that is, by an equation that implies a relation but does not, without some manipulation, state how the values of one variable are obtained from those of the other. Thus, the relation between pressure and volume given by Boyle's law is usually expressed by an equation like $p v = 100$. This relation is expressed explicitly by $p = 100/v$. Similarly, the lengths of the legs of a right triangle with hypotenuse 13 are related by the equation $x^2 + y^2 = 13^2$.

Here y can be expressed explicitly by the equation $y = \sqrt{13^2 - x^2}$. In general, implicit equations cannot be replaced by explicit ones as easily as in these examples. In fact, such a transformation is usually impossible. It is therefore necessary to deal directly with the implicit equation.

Related rate problems frequently occur in which the variables are connected by an implicit equation. Let us consider the following example.

EXAMPLE 1. A barge whose deck is 12 ft below the level of a dock is drawn toward the dock by means of a cable attached to the floor of the dock and the deck of the barge. If the cable is drawn in at the rate of 4 ft per sec, how fast is the barge approaching the dock when it is 20 ft away?

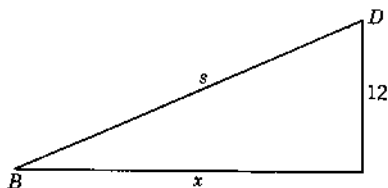


FIG. 6

SOLUTION. The situation is shown in Fig. 6, with the length of cable extending from the barge B to the dock D denoted by s , and the distance of the barge from the dock by x .

From geometry, we have

$$s^2 - x^2 = 144. \quad (13.1)$$

We look upon both x and s as functions of t such that, no matter how t varies, $s^2 - x^2$ is always equal to 144. The rate of change of $[s(t)]^2 - [x(t)]^2$ is therefore 0. Hence, differentiating, we have

$$D_t(s^2 - x^2) = 0, \quad \text{or} \quad 2sD_t s - 2xD_t x = 0.$$

From this result we find that $D_t x = \frac{s}{x} D_t s$. Now, since s is decreasing 4 ft per sec, $D_t s = -4$. At the moment when $x = 20$, we have, from (13.1), $s = \sqrt{544}$. Hence, at this moment, $D_t x = (\sqrt{544}/20)(-4) = -4.66$, approximately. Therefore, the barge is approaching the dock at approximately 4.66 ft per sec.

Since equations of curves are often given in implicit form, it is desirable to be able to find the slopes of the tangents from this form.

EXAMPLE 2. Find the slope of the tangent to the curve $x^2 + y^2 + 3xy + 5 = 0$ at the point $(-2, 3)$.

SOLUTION. Here we suppose y to be a function of x such that the equation is always satisfied no matter what value of x is taken, within some range. Then $D_x(x^2 + y^2 + 3xy + 5) = 0$, since the expression in parentheses must always have the value 0. Performing the differentiation, we have

$$D_x(x^2) + D_x(y^2) + D_x(3xy) + D_x(5) = 0.$$

Since

$$D_x(x^2) = 2x, D_x(y^2) = 2yD_xy,$$

$$D_x(3xy) = 3D_x(xy) = 3xD_{xy} + 3y,$$

and

$$D_x(5) = 0,$$

this becomes

$$2x + 2yD_{xy} + 3xD_{xy} + 3y + 0 = 0.$$

Transposing and dividing by $2y + 3x$, we have

$$D_{xy} = \frac{-2x - 3y}{2y + 3x}.$$

At the point $(-2, 3)$ we find that the slope is

$$\frac{-2 \cdot (-2) - 3 \cdot 3}{2 \cdot 3 + 3 \cdot 2} = -\frac{5}{12}.$$

EXERCISES

1. Find D_{xy} from each of the following exercises, assuming in each that the equation defines y as a function of x .

(a) $x^2 + y^2 = 9$.

(b) $xy = 12$.

(c) $x^2 - y^2 + 3y = 10$.

(d) $x^2y + xy^2 - y^3 = 1$.

(e) $\sqrt{x-1} + \sqrt{y-4} = 5$.

(f) $\frac{\sqrt{x-1}}{\sqrt{y}} = x$.

2. Find the slope of the tangent to each of the following curves at the point specified.

(a) $x^2 + y^2 - 10x + 4y + 4 = 0$, at $(1, -5)$.

(b) $y^2 - x^2 = 0$, at $(4, 8)$.

(c) $xy = 24$, at $(8, 3)$.

(d) $x^2y = 20$, at $(2, 5)$.

(e) $x^3 + y^3 - 2xy = 0$, at $(1, 1)$.

3. Solve Exercise 8, Art. 7, by differentiating the implicit equation $pv = k$.

4. The ordinate of a point describing the curve $4x^2 + 9y^2 = 288$, counterclockwise, is increasing at the rate of 3 units per sec. How rapidly is the abscissa changing when the ordinate is 4?

5. A ship proceeding due north at the rate of 18 mph passes a lighthouse 12 miles west of its course at 3 PM. How fast is the ship receding from the lighthouse at 3:30 PM?

6. A ladder 26 ft long leans against a vertical wall. How fast is the top descending if the lower end is moving horizontally away from the wall at the rate of 3 ft per sec and is now 10 ft away?

14. Summary. In this chapter we have been concerned with methods of differentiation. Although the differentiation of algebraic functions was the goal of our efforts, the general theorems are by no means restricted

to algebraic functions, and will be brought into use as we extend the processes of calculus to other functions in later chapters. We list these theorems together at this point.

$$\text{Sum formula: } D(u + v) = Du + Dv. \quad (5.2)$$

$$D(cu) = cDu. \quad (5.3)$$

$$\text{Product formula: } D(uv) = uDv + vDu. \quad (6.1)$$

$$\text{Quotient formula } D\left(\frac{u}{v}\right) = \frac{vDu - uDv}{v^2}. \quad (6.2)$$

$$\begin{aligned} \text{Chain rule: } & \text{If } y = f(u), u = \phi(x), \\ & D_x y = D_u y \cdot D_x u. \end{aligned} \quad (8.1)$$

$$\text{Derivative of an inverse: } D_x y = 1/D_y x. \quad (9.1)$$

Derivative of a function expressed parametrically:

$$\text{If } u = f(t), x = g(t), D_x y = D_t y / D_t x. \quad (12.1)$$

In addition, a method of obtaining a derivative from an implicit relation was explained and illustrated.

The one fundamental formula for differentiation, which, together with these general rules, facilitates the differentiation of algebraic functions, is the power formula and its generalization by the chain rule.

$$Dx^n = nx^{n-1}. \quad (D-1)$$

$$D_x u^n = nu^{n-1} D_x u. \quad (D-1a)$$

With the perfection of methods for the calculation of derivatives, we may consider the tangent problem and the general rate problem completely solved for algebraic functions. Further ramifications and applications of this problem will be studied in the chapters that immediately follow. Later, more advanced functions will be introduced and the operation of differentiation will be extended to them. These include trigonometric functions, logarithmic functions, and exponential functions. Throughout our work, the above formulas and theorems will be freely employed. One must not, however, lose sight of the fundamental limiting process by which they were derived.

REVIEW EXERCISES

Find the derivative of each of the functions in Exercises 1-10, with respect to the variable that appears in the expression.

1. $(1 - 3x^2)^3$.

2. $2\sqrt[3]{t}$.

3. $6x^{\frac{1}{2}} - \frac{1}{x^{\frac{1}{3}}}$.

4. $r/(4 - 9r^2)$.

5. $1/(1 - 2s)^{\frac{1}{2}}$.

6. $\sqrt{2x + 1}/(2 - x)^2$.

7. $\sqrt{(1+y^2)/(1-y^2)}$. 8. $1/(x + \sqrt{4+x^2})$. 9. $(3-4x^2)^2(2x+1)^{\frac{1}{2}}$,
 10. $\sqrt{x+1} \sqrt{x^2-1}/(3x^2-4)$.

Find $D_x y$ from each of Exercises 11-20.

11. $x = 4y^3$. 12. $x = 2y^3 - y + 1$.
 13. $x^{\frac{1}{2}} + y^{\frac{1}{2}} = 1$. 14. $x^2 y^3 - x^2 + y^2 - 4 = 0$.
 15. $3x - 4 = 4y^3 + 3y$. 16. $y/x = 2 - \sqrt[3]{x}$.
 17. $y = 9u^{\frac{1}{2}}$, $u = 3x^2 - 2$. 18. $y = \sqrt{u-1}$, $u = \sqrt{x} + 1/\sqrt{x}$.
 19. $y = 8u^{\frac{1}{2}}$, $x = u^{\frac{1}{2}}$. 20. $y = 3u^2 - 5$, $x = 4/(u-2)$.

21. Find the equations of the tangents to the following curves at the points specified:

- (a) $y = x^3 - 3x^2 + 2x - 7$; at $(2, -7)$.
 (b) $x^3 = y^2$; at $(4, -8)$.
 (c) $x = 1/y - y$; at $(\frac{3}{2}, \frac{1}{2})$.
 (d) $x = t + 1$, $y = 1 - 1/t$; at $(2, 0)$.
 (e) $x^{\frac{1}{2}} + y^{\frac{1}{2}} = 4$; at $(2\sqrt{2}, 2\sqrt{2})$.

22. If a ball is thrown straight upward at a velocity of 100 ft per sec, its height at the end of t sec is s ft, where $s = 100t - 16t^2$. Find a formula for its velocity at any time. Find the velocity at $t = 3$ and at $t = 4$. At what time is its velocity 0, and how high is it then? When does it strike the ground and with what velocity?

23. If a body moves in a straight line so that the distance s at time t is given by the equation $s = t^3 - 8t^2 + 20t - 36$, find the velocity and acceleration when $t = 3$. For what value of t is the velocity equal to 0? For what value of t is the acceleration equal to 0?

24. A spherical toy balloon is being inflated at the rate of 3 cu in. per sec. How fast is the radius increasing when it is 2 in.? How fast is the area of the surface increasing at the same instant?

25. An automobile going due north and proceeding at 60 mph passes an intersection of two straight roads. A second automobile passes the same intersection 30 sec after the first, going 40 mph due east. How fast are the two automobiles separating 3 min after the second one passed the intersection?

Second and Higher Order Derivatives

1. Differentiation of a Derived Function. We have seen how to differentiate all the familiar algebraic functions, i.e., how to obtain their derived functions. Each of these derived functions is itself an algebraic function, and hence we can differentiate it, just like the original function. Thus, if

$$f(x) = x^3, \quad Df(x) = f'(x) = 3x^2.$$

Differentiating $f'(x)$, $Df'(x) = 6x$. The quantity last obtained, which is the derived function of the derived function of x^3 , is called the *second derived function* of x^3 , and is denoted by $D^2f(x) = D^2(x^3) = 6x$. The function $Df(x) = 3x^2$ is called the *first derived function* of $f(x) = x^3$. All the derived functions found in Chapter 3 were, therefore, first derived functions.

As noted in Chapter 3, the term "derivative" is often used in the sense of "derived function." Thus, we say the *first derivative* of the function $f(x) = x^3$ is $Df(x) = 3x^2$, and the *second derivative* is $D^2f(x) = 6x$. The first derivative of this function at a point, e.g., for the value $x = 3$, is $Df(x)_{x=3} = 3 \cdot 3^2 = 27$, and the second derivative at this point is $D^2f(x)_{x=3} = 6 \cdot 3 = 18$.

Several notations for first derived functions were mentioned in Chapter 3, Art. 3. Corresponding to each of these, the notation for the second derived function and second derivative at a point are as follows.

Function	y	y	y	$f(x)$	$f(x)$	$f(x)$
First derived function	Dy or $D_x y$	y'	$\frac{dy}{dx}$	$Df(x)$	$f'(x)$	$\frac{d}{dx}f(x)$
First derivative at $x = x_0$			$\left(\frac{dy}{dx}\right)_{x=x_0}$	$Df(x)_{x=x_0}$	$f'(x_0)$	$\frac{d}{dx}f(x)_{x=x_0}$
Second derived function	D^2y or $D_x^2 y$	y''	$\frac{d^2y}{dx^2}$	$D^2f(x)$	$f''(x)$	$\frac{d^2}{dx^2}f(x)$
Second derivative at $x = x_0$			$\left(\frac{d^2y}{dx^2}\right)_{x=x_0}$	$D^2f(x)_{x=x_0}$	$f''(x_0)$	$\frac{d^2}{dx^2}f(x)_{x=x_0}$

There is, of course, no reason why we must stop with second derivatives. Each second derived function is subject to further differentiation to obtain a third derived function, or third derivative; this, in turn, leads to a fourth derived function, or fourth derivative; and so on. The notations are similar to those for the second derivative. Thus, if $y = f(x) = x^3$, then

$$Dy = y' = f'(x) = 3x^2; \quad D^2y = y'' = f''(x) = 6x;$$

$$D^3y = y''' = f'''(x) = 6; \quad D^4y = y^{(4)} = f^{(4)}(x) = 0; \text{ etc.}$$

Similarly, if $y = f(x) = x^5/10 - x^4/2 + x^3 + x - 6$, then

$$Dy = y' = f'(x) = x^4/2 - 2x^3 + 3x^2 + 1;$$

$$D^2y = y'' = f''(x) = 2x^3 - 6x^2 + 6x;$$

$$D^3y = y''' = f'''(x) = 6x^2 - 12x + 6;$$

$$D^4y = y^{(4)} = f^{(4)}(x) = 12x - 12;$$

$$D^5y = y^{(5)} = f^{(5)}(x) = 12;$$

$$D^6y = y^{(6)} = f^{(6)}(x) = 0; \text{ and all higher derivatives are 0.}$$

In connection with the D , D^2 , etc., and $\frac{d}{dx}$, $\frac{d^2}{dx^2}$, etc., notations, it should be observed that, by the definitions of the higher-order derivatives,

$$D^2y = D(Dy), \quad D^3y = D(D^2y), \text{ etc.}$$

$$\frac{d^2y}{dx^2} = \frac{d}{dx} \left(\frac{dy}{dx} \right), \quad \frac{d^3y}{dx^3} = \frac{d}{dx} \left(\frac{d^2y}{dx^2} \right), \text{ etc.}$$

The symbols D and $\frac{d}{dx}$ are not quantities; they are symbols of the operation of differentiation. They nevertheless follow a power law like the law of combination of powers of quantities, because of the way in which they have been defined.

EXAMPLE 1. Find D^2y , D^3y , and $D^{(n)}y$, if $y = 1/x$.

SOLUTION. We first find Dy , applying the power formula to $y = 1/x = x^{-1}$. Thus

$$Dy = -1 \cdot x^{-2} = -x^{-2} = -1/x^2.$$

Then

$$D^2y = D(Dy) = (-1)(-2)x^{-2-1} = 2x^{-3} = 2/x^3,$$

and

$$D^3y = D(D^2y) = (-1)(-2)(-3)x^{-3-1} = -6x^{-4} = -6/x^4.$$

We see that, at each successive differentiation, the exponent in the expression differentiated becomes a factor of the result, and that the exponent is reduced by 1. Hence,

$$\begin{aligned} D^{(n)}y &= (-1)(-2) \cdots (-n)x^{n-1} \\ &= (-1)^n n! / x^{n+1}. \end{aligned}$$

EXAMPLE 2. Find D^2y and D^3y , if $y = \frac{x}{1+x^2}$.

SOLUTION. By application of the quotient formula, we find

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

Now

$$\begin{aligned} D^2y &= D \left(\frac{1-x^2}{(1+x^2)^2} \right) = \frac{(1+x^2)^2 \cdot (-2x) - (1-x^2) \cdot 2(1+x^2) \cdot 2x}{(1+x^2)^4} \\ &= \frac{(1+x^2)(-6x+2x^3)}{(1+x^2)^4} \\ &= \frac{2x^3-6x}{(1+x^2)^3}. \end{aligned}$$

Similarly,

$$\begin{aligned} D^3y &= D \left(\frac{2x^3-6x}{(1+x^2)^3} \right) = \frac{(1+x^2)^3(6x^2-6) - (2x^3-6x) \cdot 3(1+x^2)^2 \cdot 2x}{(1+x^2)^6} \\ &= \frac{(1+x^2)^2 \cdot (-6x^4+36x^2-6)}{(1+x^2)^6} \\ &= \frac{-6x^4+36x^2-6}{(1+x^2)^4}. \end{aligned}$$

EXERCISES

Find the second derived function of each of the functions in Exercises 1-8.

1. $f(x) = \frac{1}{12}x^4 - \frac{1}{5}x^3 + \frac{1}{2}x^2 + x + 7.$

2. $f(x) = x - \frac{2}{x}.$

3. $f(x) = \frac{x}{1-x}.$

4. $f(x) = \frac{2}{1+x^2}.$

5. $f(x) = \frac{x^2}{1+x^2}.$

6. $f(x) = 2\sqrt{x}.$

7. $f(x) = \frac{x}{\sqrt{x+1}}.$

8. $f(x) = x\sqrt{x^2-1}.$

Find the second and third derived functions of the functions in Exercises 9-12.

$$9. f(x) = \frac{1}{x^2}.$$

$$10. f(x) = \sqrt[3]{x}.$$

$$11. f(x) = \frac{x+1}{x-1} = 1 + \frac{2}{x-1}.$$

$$12. f(x) = (x^2 - 1)^{\frac{1}{2}}.$$

2. Geometrical Meaning of the Second Derivative. The first derivative was found in answer to the question: What is the slope of the tangent to the graph of $y = f(x)$ in rectangular coordinates? The derived function was then seen to be an expression from which the slope could be found for any value of x . Thus, for the graph of $y = \frac{1}{6}x^3$, the slope for any

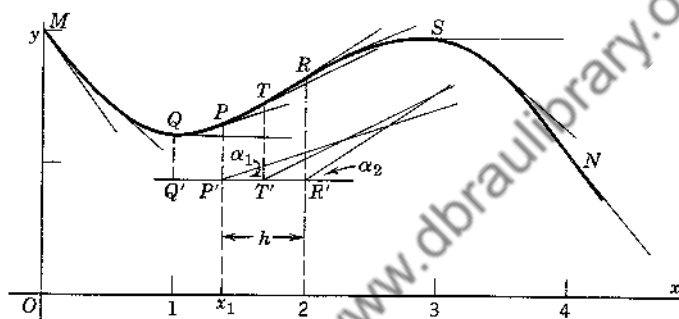


FIG. 1

value of x can be found by inserting this value of x in the derived function $\frac{1}{2}x^2$: for $x = 0$ the slope is 0, for $x = 1$ the slope is $\frac{1}{2}$, for $x = 2$ the slope is 2, etc.

In order to understand the geometrical meaning of the second derivative, we must remember that the derivative of a function of x represents the rate of change of the function with respect to x , i.e., per unit change in x . The slope of a curve, $y = f(x)$, is the rate of change of y with respect to x , or the rate of rise of the curve. The second derivative is the rate of change of the first derivative, that is, the rate of change of the slope. This rate is also with respect to x , and describes the turning of the tangent as the point of tangency moves along the curve. To get a clear geometrical picture of the meaning of the second derivative, consider the function $y = 2 - \frac{9}{5}x + \frac{6}{5}x^2 - \frac{1}{6}x^3$, represented graphically in Fig. 1. The first derivative is $y' = -\frac{9}{5} + \frac{12}{5}x - \frac{1}{2}x^2$ and the second derivative is $y'' = \frac{12}{5} - x$.

If we think of a point P moving along the curve from M to N , the tangent line at P will gradually change direction. The slope at M is $-\frac{9}{5}$ and it gradually increases to become 0 at Q , and continues to increase to

become $\frac{3}{5}$ at R , after which it gradually decreases to become 0 again at S and negative thereafter. This is shown graphically by the changing direction of the tangent as P progresses along the curve. If we consider two points P and R , and compare the slopes at these points, i.e. $\tan \alpha_1$ and $\tan \alpha_2$, we observe a change in the slope as x increases by h . The rate of change of this slope at P , with respect to x , is

$$\lim_{h \rightarrow 0} \frac{\tan \alpha_2 - \tan \alpha_1}{h} = \lim_{h \rightarrow 0} \frac{f'(x_1 + h) - f'(x_1)}{h} \\ = f''(x_1).$$

Of course, as $h \rightarrow 0$, $f'(x_1 + h) \rightarrow f'(x_1)$, just as $f(x_1 + h) \rightarrow f(x_1)$ when $h \rightarrow 0$ in obtaining the derivative $f'(x_1)$, in Chapter 3. In Fig. 1 the directions of the tangents at P , T , and R are shown on the horizontal line $Q'R'$, indicating the increase in slope as x increases from x_1 to $x_1 + h$.

Figure 2 shows three graphs of the three related functions $y = f(x)$, $y' = f'(x)$, and $y'' = f''(x)$. The function $f(x)$ is the same as that graphed in Fig. 1, viz., $f(x) = 2 - \frac{9}{5}x + \frac{6}{5}x^2 - \frac{1}{5}x^3$, and, of course, $f'(x)$ and $f''(x)$ are the first and second derived functions,

$$f'(x) = -\frac{9}{5} + \frac{12}{5}x - \frac{3}{5}x^2,$$

$$f''(x) = \frac{12}{5} - \frac{6}{5}x.$$

Some of the relationships between the graphs are pointed out on the figure.

There are two points about the geometrical interpretation of the second derivative concerning which a warning may not be superfluous. First, although the second derivative, at any point, describes the turning of the tangent line, it is not the rate of change of the angle that the tangent makes with any fixed line, e.g., the angle of inclination α . It is the rate of change of the slope, i.e., of $\tan \alpha$. Second, it is the rate of change of the slope with respect to x , and not with respect to distance measured along the curve. This is made clear by comparing the graphs of y' and y'' in Fig. 2. At a later point we shall consider a rate of change of direction that is both of the things that we have just noted that the second derivative is not, viz., the rate of change of the angle with respect to distance along the curve.

There is no necessity of seeking geometrical interpretations of derivatives beyond the second. Such interpretations are hard to grasp and of no great value in our work.

In practice, the principal use of the geometrical interpretation of the second derivative lies in the following fact, which is intuitively obvious. At any point where the second derivative is positive the right-hand end

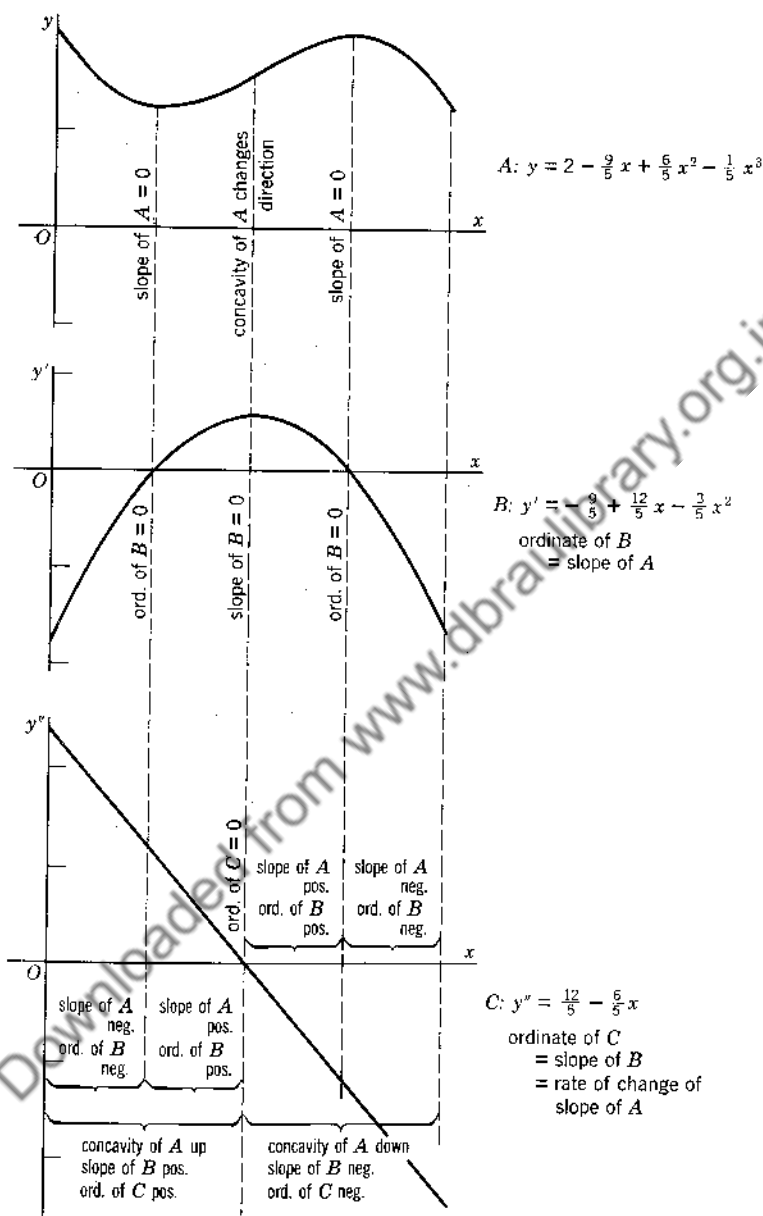


FIG. 2

of the tangent is swinging upward as we proceed from left to right along the curve; that is, the curve is concave upward, whereas at any point where the second derivative is negative the right-hand end of the tangent is swinging downward as we proceed from left to right along the curve, and the curve is concave downward. At a point where the concavity changes from upward to downward the second derivative is zero, e.g., the point R in Fig. 1. Such a point is called an *inflection point*.

The application of both the first and second derivatives to the graphs of equations will be treated further in Chapters 5 and 14.

EXAMPLE. Find the slope of the graph of $y = \frac{1}{2}x^3 - x^2 + 4x - 2$ at $x = 1$, and find the direction of concavity at this point. Sketch a small part of the curve in the neighborhood of this point.

SOLUTION. Differentiating twice, we get

$$y' = x^2/2 - 2x + 4, \quad y'' = x - 4$$

For $x = 1$ we have $y = 7/6$, $y' = 5/2$, $y'' = -1$. Hence, the curve passes through $(1, 7/6)$ with slope $5/2$, with negative rate of change of slope and, consequently, concave downward (Fig. 3).

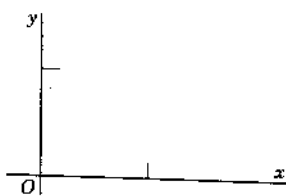


FIG. 3

EXERCISES

1. Find the slope of the graph of $y = x^3 - 4x^2 + 2x + 5$ at $x = 2$, and find the direction of concavity at this point. Sketch a small part of the curve in the neighborhood of this point.
2. Show that the graph of $y = x^2$ is always concave upward. Draw graphs of y , y' , and y'' , as in Fig. 2.
3. Show that the graph of $y = ax^2 + bx + c$ is always concave downward if $a < 0$, and always concave upward if $a > 0$.
4. Draw graphs of y , y' , and y'' , as in Fig. 2, if $y = 4 + 2x - x^2$.
5. Draw graphs of y , y' , and y'' , as in Fig. 2, if $y = x^3 - 2x^2 - 4x + 2$.

3. Rectilinear Motion. From the outset we have tried to make clear that the solution of the tangent problem carries with it the solution of a great variety of rate problems. Prominent among these is the problem of finding the velocity of a moving body. For simplicity let us consider the motion of a body, thought of as concentrated at a point P , along a straight path (Fig. 4). Let the distance of P from a fixed point O be s , and let distances toward the right be positive and those toward the left negative. If the distance s is related to the time t by a functional relation

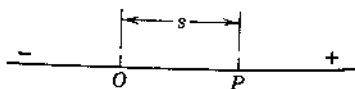


FIG. 4

$$s = f(t)$$

we know that the velocity of P is given by the derivative

$$v = f'(t).$$

If s is measured in feet and t in seconds then v is in feet per second.

Just as the point P is to the right or left of O according as s is positive or negative, and at O if $s = 0$, so if v is positive the point is moving toward the right, if v is negative it is moving toward the left, and if v is zero, it is momentarily at rest. Thus, a positive v corresponds to an algebraic increase in s , and a negative v corresponds to an algebraic decrease in s . Of course, s and v can be opposite in sign. Thus, the body may be at Q and moving to the right, in which case a negative s is increasing algebraically; it may be at R and moving to the left, in which

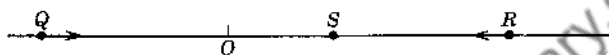


FIG. 5

case a positive s is decreasing algebraically; it may be at S and momentarily stationary (Fig. 5).

The second derivative $f''(t)$ represents the rate of change of v with respect to t and therefore the *acceleration* of the body. Thus

$$v = D_t s, \text{ and } \text{acceleration} = a = D_t v = D_t^2 s.$$

If the motion of the body is not accelerated, the velocity is constant and the motion is called *uniform* motion. The distance passed over is then simply $s = v_0 t$, where v_0 represents the constant velocity. In other words, s is a first-degree function of t . Of course, $D_t s = v_0$ and $a = D_t^2 s = 0$.

Although it may seem like a contradiction, acceleration can be a slowing down. For example, if v is positive, and a is negative, the body is moving toward the right but slowing down because of the negative acceleration. Or, if v is negative and a is positive, the body is moving toward the left but is slowing down because of the positive acceleration. But, just as the velocity is positive or negative according as s is algebraically increasing or decreasing, so the acceleration is positive or negative according as the velocity is algebraically increasing or decreasing.

4. Examples of Rectilinear Motion. As a first example let us consider the following.

EXAMPLE 1. A body (considered as concentrated at a point) moves along a straight line in such a way that its distance s (in feet) from a fixed point of the line, at time t seconds, is given by the equation $s = t^3 - \frac{15}{2}t^2 + 12t + 10$. (a) When and where is it momentarily stationary? (b) During what periods and over what distances is it moving

in a positive direction? in a negative direction? (c) During what periods is its acceleration positive? negative? (d) Show its motion by a sketch.

SOLUTION. Differentiating s ,

$$\begin{aligned} D_t s = v &= 3t^2 - 15t + 12 \\ &= 3(t-1)(t-4). \end{aligned}$$

Setting $v = 0$ to find when the body is momentarily stationary, we get

$$t = 1, \text{ or } t = 4.$$

When $t = 1$, $s = \frac{31}{2}$. Similarly, for $t = 4$, $s = 2$. Hence, the body is momentarily at rest at the end of 1 sec when it is $\frac{31}{2}$ ft, positively, from

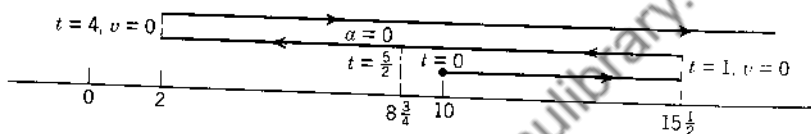


FIG. 6

the fixed point O , and again at the end of 4 sec when it is 2 ft, positively, from O .

By examination of $v = 3(t-1)(t-4)$, we see that for $t < 1$ v is positive, for t between 1 and 4 v is negative, and for $t > 4$ v is positive.

Differentiating v with respect to t , we have

$$D_t^2 s = D_t v = a = 6t - 15 = 3(2t - 5).$$

It is easy to see that a is positive for $t > \frac{5}{2}$ and negative for $t < \frac{5}{2}$. At $t = \frac{5}{2}$ the acceleration is 0. At this time the body is $8\frac{3}{4}$ positively from O . Thus, the body starts at $t = 0$, $s = 10$ with a positive velocity which is retarded by the negative acceleration until it stops at $t = 1$, $s = \frac{31}{2}$. The acceleration is still negative and produces a negative velocity which increases (negatively) until $t = \frac{5}{2}$, $s = 8\frac{3}{4}$, when the acceleration becomes momentarily 0. Thereafter the acceleration is positive. This has the effect of slowing down the negative velocity (increasing it algebraically) until it becomes 0 at $t = 4$, $s = 2$. The acceleration is still positive, producing a positive velocity which increases continually after $t = \frac{5}{2}$. The motion is shown diagrammatically in Fig. 6.

If s is graphed against t on a rectangular coordinate system we obtain the graph in Fig. 7.

The most familiar example of rectilinear motion is that of a body moving vertically under the influence of gravity. We shall assume that other forces, such as air resistance, are absent and that the force of

gravity is constant. Results obtained with these assumptions are generally fairly close approximations to the facts as long as the motion takes place near the surface of the earth at relatively low speeds. Experiments in physics have shown that near the surface of the earth the acceleration produced by the force of gravity is approximately 32 ft per sec per sec in the downward vertical direction. Since acceleration is the rate of change of velocity, it is expressed in units of velocity per unit time. Thus, if velocity is expressed in feet per second, 32 ft per sec per sec means the

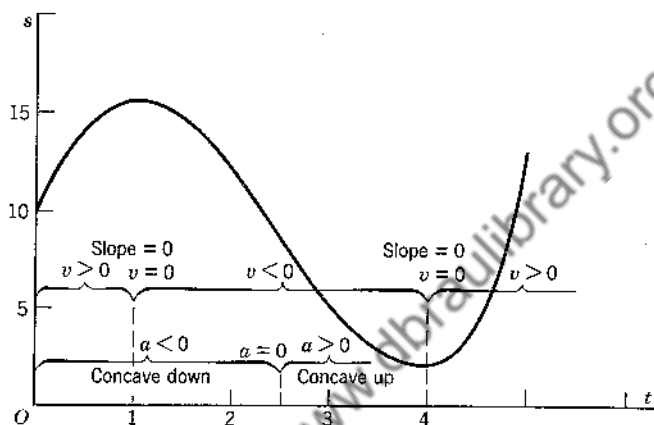


FIG. 7

rate of 32 units of velocity per unit time. For brevity, this is also expressed 32 ft per sec², or 32 ft/sec².

Since acceleration is equal to the derivative of velocity, we see that $a = 32$ would result from the variable velocity, $v = 32t + v_0$, where v_0 denotes the initial velocity, i.e., when $t = 0$. Moreover, we can see that this velocity would result from the following expression for s :

$$s = 16t^2 + v_0t + s_0,$$

where s_0 denotes the distance of the body from the point from which measurements are taken, when $t = 0$. This follows from the fact that

$$v = D_t s = D_t(16t^2 + v_0t + s_0) = 32t + v_0.$$

A more complete discussion of these relationships will appear at a later point (Chapter 6, Art. 5).

The above expressions for s , v , and a are based on the assumption that the downward direction is positive, since the force of gravity produces a downward acceleration, which has been given the positive sign (+ 32).

If we wish to consider the upward direction positive the equations should be changed to

$$a = -32, \quad (4.1)$$

$$v = -32t + v_0, \quad (4.2)$$

$$s = -16t^2 + v_0t + s_0. \quad (4.3)$$

EXAMPLE 2. A body is dropped from the top of a building 120 ft high. When will it reach the ground and with what velocity?

SOLUTION. Since it is merely dropped, its initial velocity is 0. If we think of the distances as measured from the ground, and the upward direction is positive, the initial distance is 120. Putting $v_0 = 0$, $s_0 = 120$ in equations (4.2) and (4.3), we have

$$v = -32t, \quad (4.4)$$

$$s = -16t^2 + 120. \quad (4.5)$$

To find when the body will reach the ground we set s equal to 0. Thus, from (4.5),

$$-16t^2 + 120 = 0, \quad 16t^2 = 120, \quad t^2 = 15/2,$$

and

$$t = \sqrt{15/2} = 2.7, \text{ approximately.}$$

Putting $t/\sqrt{15/2}$ in (4.4), we find

$$v = -32 \sqrt{15/2} = -86.4, \text{ approximately.}$$

Hence, the body strikes the ground in approximately 2.7 sec, with a velocity of approximately 86.4 ft per sec downward.

EXAMPLE 3. A ball was thrown upward from a building 91 ft high and struck the ground 6.5 sec later. What was the initial velocity? How high did the ball rise? With what velocity did it strike the ground?

SOLUTION. In equation (4.3) with $s_0 = 91$, we let $t = 6.5$ and $s = 0$, obtaining

$$0 = -16(6.5)^2 + v_0(6.5) + 91.$$

Solving this equation for v_0 , we get $v_0 = 90$. Thus, the initial velocity was 90 ft per sec. Putting this value in (4.4) and (4.5), we have

$$s = -16t^2 + 90t + 91, \quad (4.6)$$

$$v = -32t + 90. \quad (4.7)$$

The highest point was reached when the upward or positive velocity became 0. Immediately after this time the velocity was negative and the body was falling. Setting $v = 0$ in equation (4.7), we have

$$0 = -32t + 90, \quad \text{whence} \quad t = \frac{45}{16} = 2\frac{13}{16}.$$

The height of the ball at this time is found by putting $t = \frac{45}{16}$ in Equation (4.6),

$$s = -16\left(\frac{45}{16}\right)^2 + 90\left(\frac{45}{16}\right) + 91 = 217\frac{9}{16}.$$

The ball therefore reached a maximum height of $217\frac{9}{16}$ ft, $2\frac{13}{16}$ sec after it was thrown.

To find its velocity when it struck the ground we put $t = 6.5$ in (4.7),

$$v = -32(6.5) + 90 = -118.$$

Hence, the ball struck the ground with a velocity of 118 ft per sec downward.

EXERCISES

A particle moves along a straight line in such a way that its distance s from a fixed point at time t is given by one of the following equations. In each case describe the motion, telling (a) in what intervals the particle moves in a positive direction and in what intervals in a negative direction; (b) where the particle is when it changes the direction of its motion; (c) when the acceleration is positive and when it is negative. Consider only positive values of t .

1. $s = t^2 - 6t + 8.$

2. $s = t^3/3 - 7t^2/2 + 10t + \frac{1}{3}.$

3. $s = t^3 - 12t + 1.$

4. $s = 1/(1 + t^2).$

5. $s = 4 + t + 1/(1 + t).$

6. $s = (1 + t^2)/(1 + t).$

5. Higher Derivatives of Parametric Functions. Derived functions of order higher than the first are obtained by the same rules as those of the first order. Consequently, the methods of Chapter 3 are adequate for algebraic functions. Some directions and precautions regarding higher differentiation of functions expressed parametrically and those expressed implicitly, however, are worthy of note.

In Chapter 3, Art. 12, we found that if $y = f(u)$, $x = g(u)$, then

$$D_x y = \frac{D_u y}{D_u x} = \frac{f'(u)}{g'(u)}. \quad (5.1)$$

The result, $\frac{f'(u)}{g'(u)}$, is a function of u . Let us call it $\phi(u)$. To find the

second derivative $D_x^2 y$, we note that

$$D_x^2 y = D_x y', \quad \text{where} \quad y' = D_x y = \phi(u).$$

Thus, we have to find $D_x y'$ from the parametric equations

$$y' = \phi(u), \quad x = g(u).$$

* The primes, as usual, denote differentiation with respect to the variable that appears in the functional symbol, in this case u . y' , on the other hand, is the derivative of y with respect to x .

Applying the same process as that already used to obtain $D_x y$, we find

$$D_x y' = \frac{D_u y'}{D_u x} = \frac{\phi'(u)}{g'(u)}. \quad (5.2)$$

Now, since $\phi(u) = \frac{f'(u)}{g'(u)}$, we can express the second derivative entirely in terms of the functions $f(u)$ and $g(u)$. Using the quotient formula,

$$\phi'(u) = \frac{g'(u)f''(u) - g''(u)f'(u)}{[g'(u)]^2}.$$

Substituting this result in (5.2),

$$D_x y' = D_x^2 y = \frac{g'(u)f''(u) - g''(u)f'(u)}{[g'(u)]^3}. \quad (5.3)$$

In practice it is usually better to apply the process indicated by formula 5.2, rather than depend upon remembering formula 5.3. The process can be continued so as to obtain a derivative of any order. Example 1 illustrates this.

EXAMPLE 1. Obtain $D_x^2 y$ and $D_x^3 y$ if

$$x = u^2 + 1, \quad y = 2u - 1.$$

SOLUTION. The first derivative, from formula 5.1, is

$$D_x y = 2/2u = 1/u = \phi(u).$$

The second derivative, from formula 5.2, is

$$D_x^2 y = D_x y' = \frac{-1/u^2}{2u} = -1/2u^3.$$

The student should be warned that $D_x^2 y$ is *not* equal to $\phi'(u)$. The latter would be $D_u(D_x y)$ instead of $D_x(D_x y)$. Each differentiation will involve a division by $D(u^2 + 1)$, or $2u$. Thus

$$D_x^3 y = D_x(D_x^2 y) = \frac{D_u(-1/2u^3)}{D_u(u^2 + 1)} = \frac{3/2u^4}{2u} = \frac{3}{4u^5}.$$

EXAMPLE 2. A curve is given by the parametric equations

$$x = 2t/(1 + t^2), \quad y = (1 - t^2)/(1 + t^2).$$

Find the coordinates of the point where $t = \frac{1}{2}$, the slope at this point, and the direction of concavity.

SOLUTION. By formula 5.1, the first derivative is

$$y' = \frac{(1+t^2)(-2t) - (1-t^2)(2t)}{(1+t^2)^2} \bigg/ \frac{(1+t^2)2 - 2t \cdot 2t}{(1+t^2)^2}$$

$$= -\frac{4t}{(1+t^2)^2} \bigg/ \frac{2-2t^2}{(1+t^2)^2} = -\frac{2t}{1-t^2}.$$

Using formula 5.2, the second derivative is

$$y'' = -\frac{(1-t^2)2 - 2t(-2t)}{(1-t^2)^2} \bigg/ \frac{2-2t^2}{(1+t^2)^2}$$

$$= -(1+t^2)^3/(1-t^2)^3.$$

When $t = \frac{1}{2}$, we obtain $x = \frac{4}{5}$, $y = \frac{3}{5}$, $y' = -\frac{4}{3}$, $y'' = -\frac{125}{27}$. Since y'' is negative, the concavity is downward. The curve in the vicinity of this point is shown in Fig. 8.

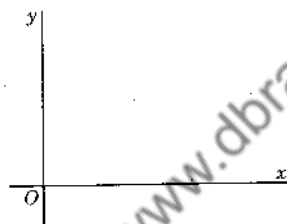


FIG. 8

EXERCISES

Find $y' = D_x y$ and $y'' = D_x^2 y$ in each of the parametric relations.

1. $x = t$, $y = 3t^2$.

2. $x = 4 + t$, $y = 1 + t^2$.

3. $x = 1/t$, $y = 12t$.

4. $x = 1 + t$, $y = 1 - t$.

5. $x = (1+t)^2$, $y = (1-t)^2$.

6. $x = \sqrt{1-u}$, $y = \sqrt{1+u}$.

7. $x = t + 1/t$, $y = 1/t$.

8. $x = \frac{3t}{1+t^2}$, $y = \frac{3t^2}{1+t^2}$.

6. Higher Derivatives of Implicit Functions. The first derivative of an implicit function was found by treating the given equation as an identity in the independent variable. The second derivative can now be found by differentiating the first. Thus, if y is a function of x defined implicitly by the equation $x^2 + y^2 = 1$, we find that $2x + 2y D_x y = 0$, and $D_x y = -x/y$. Differentiating this result by use of the quotient formula, we have

$$D_x^2 y = D_x(D_x y) = D_x \left(-\frac{x}{y} \right) = -\frac{y - x D_x y}{y^2}.$$

Inserting $-\frac{x}{y}$ for $D_x y$, and noting that $x^2 + y^2 = 1$, we obtain

$$D_x^2 y = -\frac{y - x(-x/y)}{y^2} = -\frac{y^2 + x^2}{y^3} = -\frac{1}{y^3}.$$

The second derivatives of other functions expressed implicitly can be found in the same way. The extension to higher derivatives involves no new idea.

No special formulas are needed for second and higher derivatives of inverse functions. Functions expressed this way can be treated as implicit relations. Indeed, it is often better to treat them so from the start. An example will make this clear.

EXAMPLE. If $x = \sqrt{1 + y^3}$, find $D_x^2 y$.

SOLUTION. We may write $x^2 = 1 + y^3$ and differentiate implicitly, obtaining $2x = 3y^2 D_x y$. Hence,

$$D_x y = \frac{2x}{3y^2}.$$

Now,

$$\begin{aligned} D_x^2 y &= D_x(D_x y) = \frac{3y^2 \cdot 2 - 2x \cdot 6y \cdot D_x y}{9y^4} \\ &= \frac{6y^2 - 12xy(2x/3y^2)}{9y^4} \\ &= \frac{18y^4 - 24x^2 y}{27y^5} \\ &= \frac{6y^3 - 8x^2}{9y^5}. \end{aligned}$$

EXERCISES

Find $D_x^2 y$ in each of the following.

1. $x^2 - y^2 = 1$.

2. $xy = 24$.

3. $9x^2 + 25y^2 = 225$.

4. $9y^2 - 4x^2 = 36$.

5. $x^2 = (3 - y)/y$.

6. $y^3 x = 1 + x$.

Maxima and Minima

1. The Problem of a Maximum or a Minimum. In our work with rates we have encountered functions that increased in part of the range and decreased in another part. Let us recall Example 1, Art. 4, Chapter 4. At first the velocity is positive, that is, the distance s increases with the time. Then, for a time, the velocity is negative and s decreases, and later the velocity becomes positive again and s increases. If the velocity changes gradually, as it does in this example, the positive velocity which the body has at first must slow up, become momentarily zero, and then negative. This negative velocity must eventually slow up to 0, thereafter becoming and remaining positive. Thus, the ranges of positive and negative velocities are separated one from another by points of zero velocity. The first such point in this example is the point of greatest advance of the body toward the right for a time (not forever), and the second of these points is the point of greatest distance to the left for a time (actually the farthest to the left that it ever goes in this case). Another way to describe these points is to say that the first is a point of maximum value of s and the second is a point of minimum value of s . Here "maximum" means maximum in comparison with values of s just before and just after this value is reached. In the present example we found this "relative" maximum value of s to be $\frac{31}{2}$, reached when $t = 1$, and we found the minimum value of s to be 2, reached when $t = 4$.

The maximum and minimum problems associated with this example and described above were solved in Chapter 4, although we did not emphasize this part of the example. In order to focus attention on these problems here, we describe them as follows.

We have s as a function of t , $s = f(t)$. We are to find a value of t that makes s a maximum or a minimum relative to other values of s for nearby values of t . We have to find, also, what these maximum and minimum values of s are. In the example cited above, the values called for by this problem were found during the discussion in Chapter 4. They were the values of t and s corresponding to $v = 0$, that is, to $D_t s = 0$. This suggests a simple method of finding values of the independent variable

x making a function $f(x)$ maximum or minimum, namely, find $f'(x)$, set it equal to 0, and solve the resulting equation for x .

In a large number of practical cases this method will yield the results we seek. There are exceptions, which will be noted, but they do not often occur in concrete, practical problems. A value of x found by solving the equation $f'(x) = 0$ may correspond either to a maximum or to a minimum. If it is not obvious which one, this will be a matter for further investigation. We shall see, also, that such a value may produce neither a maximum nor a minimum. In practical cases, however, it is a relatively simple matter to classify values of x so found. Our principal problem is to find them.

The term "maximum-minimum problem" may refer to a problem of finding a value of the independent variable that produces a maximum or minimum of a function equal to the dependent variable, finding this maximum or minimum value, or both.

The importance of the maximum-minimum problem can be appreciated to some extent from a mere statement of some examples.

1. Rectangular packing cases are to be constructed, each with the base twice as long as it is wide and all of them holding the same given volume. What should the dimensions be in order that the least amount of material will be used?

2. The strength of a beam of a given material, having a fixed length and a rectangular cross section, varies directly as the product of the breadth and the square of the depth. Find the dimensions of the strongest beam that can be sawed from a log whose cross section is a circle 6 in. in radius.

3. Two sources of light of intensities a and b , respectively, are d feet apart. If the intensity of illumination at any point due to any source is directly proportional to the intensity of the source and inversely proportional to the square of the distance from the source, find the point on the line joining the two given sources at which the illumination is least.

4. The cost of fuel for running a train is proportional to the square of the speed maintained and is \$50 per hr when the speed is 20 mph. Other charges, such as labor, for example, are \$200 per hr. What should be the speed for a 500-mile trip in order that the cost shall be as little as possible?

2. Maxima and Minima and the Tangent Problem. Since we have considered the problem of the slope of the tangent to a curve as the prototype of all rate problems, let us examine the relationship of the tangent problem to the problem of maxima and minima.

EXAMPLE 1. As a first example, we consider the graph of

$$y = -20 + 10x - x^2.$$

From analytical geometry we know that the curve is a parabola with its axis vertical and concave downward. It must therefore have a maximum point at its vertex, which is found by methods of analytical geometry to be (5, 5) (Fig. 1).

In general, we do not have formulas for maximum or minimum points, as we have in this example. How, then, shall we find them? We observe that the slope of the tangent is positive for values of x to the left of the maximum point and negative for values of x to the right of this point.

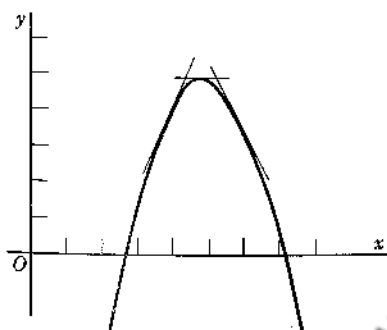


FIG. 1

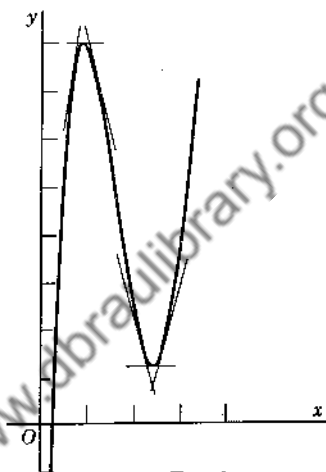


FIG. 2

It is clearly equal to 0 at the maximum point. We therefore find the derivative

$$Dy = 10 - 2x,$$

and set this equal to 0. Thus, for the value of x at the maximum point

$$10 - 2x = 0,$$

and

$$x = 5.$$

When $x = 5$, $y = 5$, and we conclude that the maximum point is (5, 5), as before.

EXAMPLE 2. As a second example, let us consider the graph of

$$y = 4x^3 - 21x^2 + 30x - 5.$$

Using a short table of values of x and y , we obtain the graph shown in Fig. 2.

x	-1	0	1	2	3	4
y	-60	-5	8	3	4	35

In this example there is neither a highest nor a lowest point. There is, however, a point of the curve that is higher than any other point of the curve in its immediate neighborhood, and another point that is lower than any other point of the curve in its immediate neighborhood. These points are called *relative* maximum and minimum, respectively. To find the abscissas of these points we again make use of the fact that at each of them the slope of the tangent is equal to 0. Differentiating, we get

$$Dy = 12x^2 - 42x + 30.$$

Setting this equal to 0 to find the values of x for which the slope is 0,

$$12x^2 - 42x + 30 = 0.$$

The roots of this equation are 1 and $\frac{5}{2}$. Since $y = 8$ when $x = 1$, and $y = \frac{5}{4}$ when $x = \frac{5}{2}$, we conclude that the point $(1, 8)$ is a relative maximum and $(\frac{5}{2}, \frac{5}{4})$ is a relative minimum.

In deciding that $(1, 8)$ is a maximum point and that $(\frac{5}{2}, \frac{5}{4})$ is a minimum point we are, of course, assisted by our knowledge of the

general course of the curve. If we know very little about the form of a curve we may require tests to determine whether a point at which the slope is zero is a maximum or a minimum or, indeed, neither one (see Fig. 3 above). Such tests will be described at a later point when we shall consider more complicated curves.

The fact that more discussion is needed to dispose of maximum and minimum points completely can be surmised from examining the curves in Figs. 3-7. Figure 3 shows the graph of the equation $y = x^3$. Here, $Dy = 3x^2$, and the slope = 0 for $x = 0$. But this is neither a maximum point nor a minimum point.

Figure 4 shows the graph of the equation

$$y = x^3/27 + 4x + 4.$$

Here

$$Dy = x^2/9 + 4.$$

Setting this derivative equal to 0, we have

$$x^2/9 + 4 = 0$$

which is satisfied only by the imaginary roots $6i$ and $-6i$. Thus, the

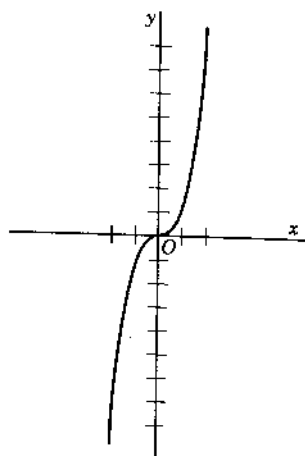


FIG. 3

derivative is not equal to 0 for any real value of x and there is no maximum or minimum point on the curve.

Figures 5 and 6 show curves at the maximum points of which the slope is not equal to 0. These are not smooth curves, however, that is, curves with gradually turning tangents. Their equations are more complicated than any we are likely to encounter in applications, and we shall not

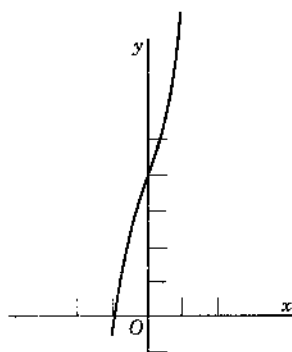


FIG. 4

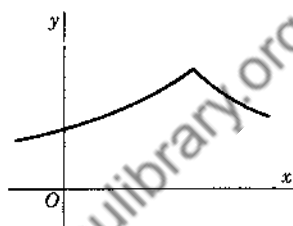


FIG. 5

consider them here. Each of their derived functions is discontinuous at the maximum point. Figure 7 shows a curve having its maximum point at an end-point, where the slope is not 0.

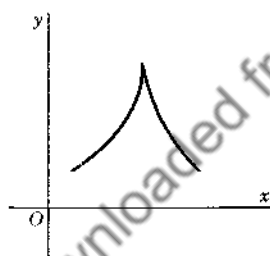


FIG. 6

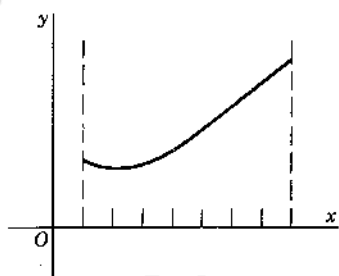


FIG. 7

3. Definitions and Theorems. In order to avoid continued dependence on geometric intuition and to make the ideas of maxima and minima precise, we make the following definitions.

DEFINITION 1. A function $f(x)$ is said to be *increasing* at $x = x_1$ if, for sufficiently small positive h , $f(x_1 - h) < f(x_1) < f(x_1 + h)$. A function is said to be *decreasing* at $x = x_1$ if, for sufficiently small positive h , $f(x_1 - h) > f(x_1) > f(x_1 + h)$.

DEFINITION 2. A point $x = x_1$ corresponds to a relative maximum of a function $f(x)$ if $f(x_1) - f(x) > 0$ for every x in the interval of definition

of $f(x)$ that lies in some neighborhood of x_1 . If x_1 is in the interior of the interval of definition of $f(x)$, i.e., it is not an end-point, this means that, for sufficiently small positive h , $f(x_1) - f(x) > 0$ for every x in the interval $x_1 - h < x < x_1 + h$.

A relative minimum is defined in the same way as a relative maximum except that the inequality $f(x_1) - f(x) > 0$ is replaced by $f(x_1) - f(x) < 0$.

At a maximum or a minimum of a function in the interior of the interval of definition of the function, the function is clearly neither increasing nor decreasing.

We bring the derivative into this discussion by means of the following theorems, in which we assume that x_1 is in the interior of the interval of definition of the function.

THEOREM 1. If $f'(x_1) > 0$, then $f(x)$ is an increasing function at $x = x_1$.

THEOREM 2. If $f'(x) < 0$, then $f(x)$ is a decreasing function at $x = x_1$.

THEOREM 3. If $f(x)$ has a relative maximum or a relative minimum at $x = x_1$, and possesses a unique derivative there, then $f'(x_1) = 0$.

We shall prove Theorem 1. The proof of Theorem 2 is similar. Theorem 3 will then be seen to follow. By the hypothesis of Theorem 1, $f'(x_1) > 0$. Since

$$\lim_{h \rightarrow 0} \frac{f(x_1 + h) - f(x_1)}{h} = f'(x_1),$$

we have

$$\frac{f(x_1 + h) - f(x_1)}{h} = f'(x_1) + \varepsilon(h),$$

where

$$\varepsilon(h) \rightarrow 0 \text{ as } h \rightarrow 0.$$

Hence, for sufficiently small h ,

$$\frac{f(x_1 + h) - f(x_1)}{h} > 0.$$

If h is positive, it follows that $f(x_1 + h) - f(x_1)$ is positive, and $f(x_1) < f(x_1 + h)$. If h is negative, say $h = -k$, we have

$$\frac{f(x_1 - k) - f(x_1)}{-k} > 0.$$

Multiplying by $-k$, we obtain $f(x_1 - k) - f(x_1) < 0$, and $f(x_1 - k) < f(x_1)$. Thus, the condition that $f(x)$ be an increasing function is satisfied.

In order to prove Theorem 3 we note that if $f(x)$ has a maximum at $x = x_1$, the function is neither increasing nor decreasing at that point.

Hence, $f'(x_1)$ is neither greater than 0 nor less than 0. Therefore, $f''(x_1) = 0$. The proof is the same for a minimum.

It is especially important to note that Theorem 3 applies in the interior of an interval of definition of a function that has a unique derivative at the point in question. Figures 5, 6, and 7 illustrate maxima where the conditions for the theorem are not satisfied.

The converse of Theorem 3 is not true, as shown by the function $y = x^3$, whose graph appears in Fig. 3.

Points at which the derivative is equal to 0 are sometimes called *stationary points*. They are points at which, if a particle were progressing along a curve from left to right, its vertical motion would be momentarily stopped. At a maximum point the particle has stopped rising and is about to descend. At a minimum point the particle has stopped going downward and is about to go upward. At a point like (0, 0) in Fig. 3, the particle has momentarily stopped its rise but is about to start upward again.

At the point (0, 0) in Fig. 3 the direction of concavity of the curve changes, and, hence, this point is an inflection point with $D^2y = 0$ (Chapter 4, Art. 2). A point like this one, which is an inflection point with a horizontal tangent, is often called a *level inflection point*. Maximum and minimum points at which the slope is 0 are also often called *level points*.

Stationary points are frequently called *critical points*, and the value of the independent variable at such a point is called a *critical value*. In the study of complicated curves it is often a considerable task to determine the nature of each critical point, i.e., whether it is a maximum, a minimum, or a level inflection point. Later we shall derive tests for this determination. With relatively simple curves and in practical applications, however, there is seldom any difficulty of this sort.

EXAMPLE. Examine the graph of

$$y = 3x^4 - 20x^3 + 42x^2 - 36x + 12$$

for maxima, minima, and points of inflection, and draw the graph.

SOLUTION. Differentiating, we have

$$Dy = 12x^3 - 60x^2 + 84x - 36 = 12(x - 1)^2(x - 3),$$

and

$$D^2y = 12(3x^2 - 10x + 7) = 12(x - 1)(3x - 7).$$

Setting $Dy = 0$, we find stationary points at $x = 1$ and $x = 3$. Substituting these values in the expression for y , the coordinates of the stationary points are found to be (1, 1) and (3, -15). Setting $D^2y = 0$, we find as possible inflection points the points (1, 1) and $(\frac{7}{3}, -\frac{229}{27})$.

In order to sketch the graph, we note that for numerically large values of x , either positive or negative, y is large positively. Moreover, the only points of level tangency are $(1, 1)$ and $(3, -15)$. Hence, the curve must be above the point $(1, 1)$ to the left of this point and descend toward it. It must continue to descend between $(1, 1)$ and $(3, -15)$, and thereafter rise. It is easily verified that $Dy < 0$ everywhere to the left of $x = 3$

except at $x = 1$, and hence the function decreases in this range. To the right of $x = 3$, $Dy > 0$ and the function is increasing. The graph is shown in Fig. 8. The point $(1, 1)$ is obviously a level inflection point and $(3, -15)$ is a minimum point. There must be an inflection point between $(1, 1)$ and $(3, -15)$ in order for the direction of concavity to change. This is the point $(\frac{7}{3}, -\frac{229}{27})$.

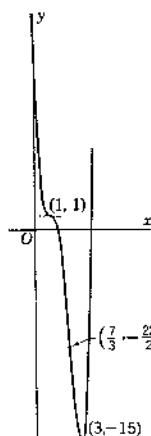


FIG. 8

EXERCISES

1. Find a maximum point of the graph of each of the following.

(a) $y = 2 - x^2$.

(b) $y = 4 + 4x + x^2$.

(c) $y = 2 - 10x - x^2$.

(d) $y = -16x + x^2 + 2$.

2. Find a minimum point of the graph of each of the following:

(a) $y = x^2 - 2$.

(b) $y = 4 + 4x + x^2$.

(c) $y = x^2 + 5x + 6$.

(d) $y = 3x^4 - 4x^3 - 6x^2 + 24x + 10$.

3. Find the level points of the graph of each of the following, determine from the graph which are maxima, which are minima, and which are inflection points, and draw the graph.

(a) $y = 3 - x^3$.

(b) $y = 2x^3 - 9x^2 + 12x - 6$.

(c) $y = x^3 + 3x^2 - 9x - 8$.

(d) $y = x^4 - 2x^2 - 2$.

(e) $y = 8x^3 - x^4$.

(f) $y = 3x^5 - 5x^3 + 15$.

4. At an inflection point $D^2y = 0$, whereas the curve is concave up if $D^2y > 0$ and concave down if $D^2y < 0$. From these considerations, formulate tests to determine the nature of a critical point. Apply these tests in Exercise 3. [N.B. If $D^2y = 0$, the point is not necessarily an inflection point; e.g., $y = x^4$ at $(0, 0)$.]

5. Show that $y = 8/(x^2 + 4)$ has a maximum point at $(0, 2)$, and two inflection points symmetrically located with regard to the y -axis. Draw the curve.

6. Find maximum and minimum points and inflection points of the curve $y = (x^2 + 8)/(x^2 + 4)$, and sketch the curve.

7. Find maximum and minimum points of the curve $y = x^2/(x - 3)$, and show that there are no inflection points. Draw the curve.

8. Solve Exercise 7 after replacing the given equation by $y = x + 1/x$.

9. Find maximum and minimum points of the curve $y = 4/x + 9/(2 - x)$ and sketch the curve.

10. Show that if $y = ax^3 + bx^2 + cx + d$ and $Dy = 0$ for two values of x , say x_1 and x_2 with $x_1 < x_2$, then x_1 corresponds to a maximum of y , and x_2 to a minimum, provided $a > 0$; x_1 corresponds to a minimum, and x_2 to a maximum, if $a < 0$. [Hint: y is large for large values of x and large negative values for large negative values of x .]

11. If the supply y of a product is related to its price x by the equation $y = a\sqrt{x - b} + c$, in which a , b , and c are positive constants and $x > b$, show that the supply function is always increasing but that the rate of increase is a decreasing function.

4. Applications. We now return to the concrete problems proposed in Art. 1. The method of determining maxima and minima in these problems will be the same as that applied to curves in Art. 2 and Art. 3. In each problem, if we chose to do so, we could represent the function involved by a curve, and a maximum or minimum point of the curve would correspond to a maximum or minimum of the concrete function.

In the majority of practical problems, either a maximum or a minimum is to be expected, and not both. There is no difficulty in deciding whether a result corresponds to a maximum value or to a minimum value since we know in advance which to expect. The whole problem is that of determining what it is, precisely.

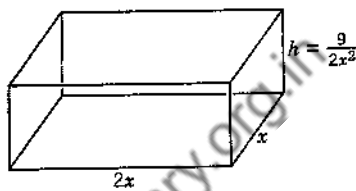


FIG. 9

EXAMPLE 1. A company contracts to manufacture packing cases, each with a rectangular base twice as long as it is wide and holding 9 cu ft. What should be the dimensions in order that the least amount of material shall be used, disregarding waste and overlapping?

SOLUTION. Under the given conditions, the amount of material needed for one packing case is proportional to its total area. We have to find the dimensions corresponding to a minimum area with a given volume. First, we try to express the area as a function of one variable, say the width x . There must be enough information given so that this can be done. Otherwise we cannot solve the problem. Thus, referring to Fig. 9, let x = width in ft, and $2x$ = length in ft. Then $2x \cdot x = 2x^2$ = area of the base = area of the top. In order to express the area of the sides and ends we need to use the height h . The area of each end is xh and that of each side is $2xh$. Hence, the whole area = $A = 2x^2 + 2x^2 + xh + xh + 2xh + 2xh$, or $A = 4x^2 + 6xh$. We now use the fact that the volume is 9 to eliminate h from this expression. Since $V = 2x \cdot x \cdot h = 2x^2h$, we have $9 = 2x^2h$, and $h = 9/2x^2$. Hence, substituting for h in the expression for A ,

$$A = 4x^2 + 27/x.$$

The range of x in this functional relation is $x > 0$, since only positive values of x have meaning in the problem. There is no upper bound to values of x .

We can now proceed as in Art. 3. Let us note the meaning of each step in this process. Differentiating,

$$DA = 8x - 27/x^2.$$

This gives us a formula for the rate of change of A per unit change in x . Setting $DA = 0$, we have

$$8x - 27/x^2 = 0.$$

We do this in order to find a critical value. In this case this is the value of x where A ceases to decrease and starts to increase (see discussion below). Solving for this critical value,

$$8x^3 = 27, \quad 2x = 3, \quad \text{and} \quad x = \frac{3}{2}.$$

We conclude that the minimum area occurs if the width is $\frac{3}{2}$ ft, the length is $2 \cdot (\frac{3}{2}) = 3$ ft, and the height is $\frac{9}{2 \cdot (\frac{3}{2})^2} = 2$ ft (since $h = 9/2x^2$).

We can see that this critical value, $x = \frac{3}{2}$, produces a minimum as follows. If x is small, A is large; for example, if $x = .1$, $A = 4(.01) + 27/.1 = 270.04$. A small increase in x from .1 produces a decrease in A . Thus, if we increase x to .2, A becomes $4(.04) + 27/.2 = 135.16$. This decrease cannot continue indefinitely as x increases to a large value. For, suppose $x = 10$; then $A = 4(100) + 27/10 = 402.7$. Hence, as x increases from small values to large values, A decreases at first and later increases, and consequently a minimum is to be expected. We can be sure that our expectation will be fulfilled, since both the function for A and the derived function are continuous except at $x = 0$, which has no meaning in the problem.

EXAMPLE 2. The greatest load which can be borne by a beam of a given material, of fixed length, and rectangular cross section, varies directly as the product of the breadth and the square of the depth. Find the dimensions of the strongest beam that can be sawed from a log whose cross section is a circle with radius 6 in.

SOLUTION. If x = breadth of the beam in inches, y = depth of the beam in inches, and S = safe load, we are told that $S = kxy^2$, where k is a constant. Here, the function to be maximized contains two variables x and y . There is some information in the statement of the problem, however, that will make it possible to eliminate one of these variables. Since the beam is to be sawed from a circular log of diameter 12 in., x and y must be legs of a right triangle having its hypotenuse equal to 12 (Fig. 10). Thus

$$x^2 + y^2 = 144$$

and

$$y^2 = 144 - x^2. \quad (4.1)$$

Substituting $144 - x^2$ for y^2 in the expression for S , we have

$$S = kx(144 - x^2),$$

or

$$S = 144kx - kx^3.$$

In this function, x must, of course, be positive and less than 12, i.e., $0 < x < 12$.

Proceeding as in Example 1, we find

$$DS = 144k - 3kx^2.$$

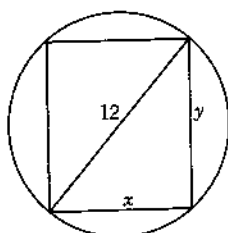


FIG. 10

Setting this derived function equal to 0,

$$144k - 3kx^2 = 0, \quad x^2 = 48, \quad \text{and} \quad x = 4\sqrt{3}.$$

(The negative value $-4\sqrt{3}$ satisfies the equation, but it does not apply to the problem.)

Inserting this value of x in (4.1), we find

$$y^2 = 144 - 48 = 96, \quad \text{and} \quad y = 4\sqrt{6}.$$

Thus, the dimensions of the beam of maximum strength are $4\sqrt{3}$ in. by $4\sqrt{6}$ in.

As in Example 1, we can see in advance whether a maximum or a minimum is to occur. It is clear that S is small if x is near 0 and increases as x increases. But if x is near 12, S is again very small, so that S must increase with x up to a point, after which it decreases as x increases.

The factor k is determined by the particular kind of wood used, the fixed length, and the units in which x , y , and S are measured. Since k does not affect the solution, it is clear that it could be changed without altering the results. Hence the results found apply to a beam of any length and any wood (supposed to be homogeneous) and are independent of the units employed to measure S . The results are independent, also, of the units employed for x and y as long as they are the same and agree with the unit of measure of the diameter. If the diameter is $12a$, where

a is any factor, then the beam of greatest strength must be $4\sqrt{3}a$ by $4\sqrt{6}a$. That is, the depth is $\sqrt{2}$ times the breadth, regardless of the size of the log.

EXAMPLE 3. Two point-sources of light of intensities a and b , respectively, are d feet apart. If the intensity of illumination at any point due to a given point source is directly proportional to the intensity of the source and inversely proportional to the square of the distance from the source, find the point on the line joining the two given sources at which the illumination is least.

SOLUTION. Let A be the source of intensity a , Fig. 11. Then, at any point P on the line AB , the intensity of illumination due to A is ka/x^2 ,

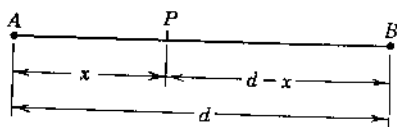


FIG. 11

and that due to B is $kb/(d-x)^2$. If I stands for the total illumination of P due to both sources,

$$I = \frac{ka}{x^2} + \frac{kb}{(d-x)^2}. \quad (4.2)$$

In this function

$$0 < x < d.$$

It is clear that at a point very near either source the value of I is great and decreases as the point recedes. This is because of the fact that if P is near either source one or the other of the denominators in (4.2) is small and the corresponding fraction is large. Since I decreases as P moves away from a position near either source it must reach a minimum somewhere between them. Hence, we anticipate a minimum, and it is our problem to discover just where it is. Differentiating (4.2),

$$D_x I = k \left(\frac{-2a}{x^3} + \frac{2b}{(d-x)^3} \right).$$

Setting this derived function equal to 0 to find the value of x producing the minimum,

$$k \left(\frac{-2a}{x^3} + \frac{2b}{(d-x)^3} \right) = 0,$$

or

$$\frac{b}{(d-x)^3} = \frac{a}{x^3}.$$

Clearing fractions,

$$bx^3 = a(d-x)^3.$$

Taking the cube roots of both sides,

$$\sqrt[3]{b}x = \sqrt[3]{a}(d-x).$$

Multiplying out and transposing,

$$x(\sqrt[3]{a} + \sqrt[3]{b}) = d\sqrt[3]{a}.$$

Hence

$$x = \frac{d\sqrt[3]{a}}{\sqrt[3]{a} + \sqrt[3]{b}},$$

and

$$d-x = d - \frac{d\sqrt[3]{a}}{\sqrt[3]{a} + \sqrt[3]{b}} = \frac{d\sqrt[3]{b}}{\sqrt[3]{a} + \sqrt[3]{b}}.$$

The student can show that the ratio of the distances $x : (d-x)$, for the minimum, is $\sqrt[3]{a} : \sqrt[3]{b}$.

EXAMPLE 4. The cost of fuel for running a train is proportional to the square of the speed maintained, and is \$50 per hr when the speed is 20 mph. Other charges, such as labor, for example, are \$200 per hr. What should be the speed for a 500-mile trip in order that the total cost shall be as little as possible?

SOLUTION. First we express the cost per hour for fuel. If C represents this cost,

$$C = ks^2,$$

where

s = speed in mph.

We now evaluate k as in any problem in variation. Since $C = 50$ when $s = 20$,

$$50 = 400k,$$

and

$$k = 50/400 = 1/8.$$

The cost per hour for fuel is therefore $\frac{1}{8}s^2$. Since other costs are \$200 per hr, the total cost per hour is $\frac{1}{8}s^2 + 200$. Now, the number of hours needed to make a 500-mile trip at s miles per hour is $500/s$. Therefore, the total cost for the trip, A dollars, is

$$A = \frac{500}{s} \left(\frac{1}{8}s^2 + 200 \right) = \frac{125}{2}s + \frac{100,000}{s},$$

in which $s > 0$. Differentiating,

$$DA = \frac{125}{2} - \frac{100,000}{s^2}.$$

Setting this quantity equal to 0 in order to find the value of s producing the minimum,

$$\frac{125}{2} - \frac{100,000}{s^2} = 0,$$

whence

$$125s^2 = 200,000, \quad s^2 = 1600, \quad \text{and} \quad s = 40.$$

Therefore, for the least total cost, the speed should be 40 mph. The student can show that the value of the function representing the cost is a minimum for $s = 40$ by reasoning as before in Examples 1-3, considering the values of A for very small and very large values of s .

The length of the trip is immaterial in this problem, since if we take it as L the same result is obtained. The student should do this. Common sense would lead to the same conclusion, since if 40 mph is the most economical speed for 500 miles, so it is for any number of times 500 miles or any part of 500 miles.

The method of solving problems in maxima and minima can be summarized as follows. (1) Express the variable to be maximized, or minimized, as a function of one independent variable. If more than one independent variable is present, find relations among them by which all can be expressed in terms of one of them; (2) differentiate this function with respect to the independent variable; (3) set the derived function equal to 0; (4) solve the resulting equation. Values of the independent variable thus found include the desired maximizing, or minimizing value; and (5) put the maximizing, or minimizing, value in the relation expressed in step 1, to find the maximum, or minimum, value of the dependent variable.

A variation of the above method that is often more convenient will be explained in Art. 5.

EXERCISES

1. Find the dimensions of the largest rectangle that has its perimeter equal to 64 in.
2. Find the dimensions of the rectangle of smallest perimeter that has its area equal to 400 sq in.
3. Find two positive numbers having the least sum if the product is 100. Do the same if the product is k .
4. The product of the concentration of hydrogen ions, C_{H^+} , and hydroxyl ions, C_{OH^-} , in any aqueous solution is a constant $= 10^{-14}$. At what concentration of hydrogen ions will the sum of hydrogen ions and hydroxyl ions be a minimum?
5. Find two numbers whose sum is 15, such that the product of one by the square of the other is a maximum.
6. Find the positive number such that the sum of the number and its reciprocal is least.
7. Find the number that exceeds its square by the greatest amount.
8. A rectangular garden is to be surrounded by a fence and divided into two equal

parts by a fence parallel to one side. If the area is to be 600 sq ft, what should be the dimensions in order to require the least amount of fencing?

9. Find the dimensions of the largest rectangle that can be inscribed in a circle of radius 6 in.

10. An open rectangular box with a square base is to be made from a piece of cardboard 18 in. on a side by cutting out equal squares from the corners and turning up the sides (Fig. 12). Find the dimensions of the box of maximum capacity that can be constructed in this way.

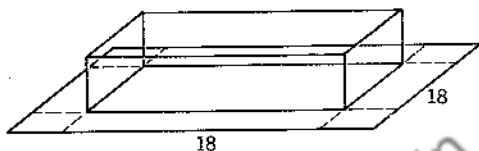


FIG. 12

11. A closed box, whose base is a rectangle twice as long as it is wide, is to have a total surface area of 108 sq in. What is the greatest capacity the box can have?

12. One side of a farm is bounded by a straight railroad. The farmer wishes to fence off a rectangular field of 20 acres, letting the fence along the railroad track serve as one boundary. What should be the dimensions of the rectangle requiring the least length of new fence, and how much fencing is needed (1 acre = 160 sq rods)?

13. Postal regulations limit the size of a package that can be sent by parcel post, by requiring that the sum of the length and the girth be not more than 84 in. Find the dimensions of the package of greatest volume with a square cross section that can be mailed. Find the volume of this package.

14. Using the limitation given in Exercise 13, find the dimensions and volume of the largest package with a circular cross section that can be mailed.

15. A rectangular box with open top and square base is to have a capacity of 32 cu ft. At 5 cents per sq ft for a metal lining, what should be the dimensions in order that the cost of lining shall be as little as possible?

16. If, in Exercise 15, it is desired to line the sides with zinc at 8 cents per sq ft, and the bottom with copper at 16 cents per sq ft, what should be the dimensions for a minimum cost?

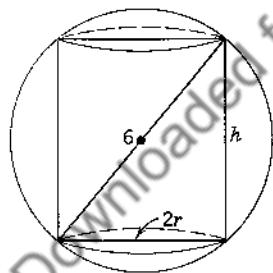


FIG. 13

17. A box in the form of a rectangular parallelepiped with square ends is to be divided into two equal compartments by a partition parallel to one side. If the box is to contain 2592 cu in., find the dimensions to require the least amount of lumber, assuming all parts to have the same thickness.

18. Find the ratio of the height to the diameter of a cylindrical tin can in order that the least amount of tin shall be used for a given capacity.

19. Find the dimensions and volume of the largest right circular cylinder that can be inscribed in a sphere 6 in. in diameter (Fig. 13).

20. A window is to have the shape of a rectangle surmounted by an equilateral triangle. If the perimeter is fixed, find the ratio of the height to the width of the rectangle in order that the greatest amount of light shall be admitted, i.e., the window shall have the maximum area.

21. If, in Exercise 20, the window has the shape of a rectangle surmounted by a semicircle, find the ratio of the height of the rectangle to its width in order to admit the greatest amount of light.

22. Two cities A and B are connected by a road that runs north from A for 12 miles and then east for 5 miles to B . This road needs repairs that will cost \$45,000 per mile. The cost of a new road would be \$75,000 per mile. It is therefore decided to repair the present road for a distance north from A and build a new road, in a northeasterly direction from the part repaired, to B . How much of the old road should be repaired in order to keep the cost of the project to a minimum, and how much will be the whole cost?

23. A camper is situated on an island A , 6 miles from the nearest point B of a straight shore. He wishes to reach a town C , 15 miles along the shore from B , in the least possible time. If he can row his boat 4 mph and walk 5 mph, where should he land?

24. The current I , from a battery, is given by $I = E/(r + R)$, in which E and r are constants representing the electromotive force and the internal resistance, respectively, and R is the variable resistance in the external circuit. The energy utilized is $I^2 R$. Show that this is greatest if $R = r$.

25. The speed of a certain autocatalytic reaction is given by the equation $s = kx(a - x)$, where x is the amount of material decomposed at any time, a is the amount of material originally present, and k is a constant. Show that the speed of the reaction is greatest when one-half the original material is decomposed.

26. The cost of erecting an office building is \$80,000 for the first floor, \$95,000 for the second, \$110,000 for the third, and so on, the cost per floor increasing in arithmetic progression as the number of floors increases. Fixed costs, such as the cost for the site, for example, amount to \$500,000. If the yearly rental is to be \$12,000 per floor, what should be the number of floors in order that the ratio of the annual income to the investment shall be greatest?

DISCUSSION. The methods of this chapter require some modification in order to apply to a problem of this kind, since the independent variable that represents the number of floors does not assume all values in an interval, but only discrete integral values. Differentiation of such a function is impossible. For convenience we may replace the numerical data of the problem by letters: A = cost of first floor, d = common difference, M = fixed costs, S = yearly rental, and n = number of floors. The expression for the ratio of income to investment is then

$$I = f(n) = \frac{Sn}{dn^2/2 + (A - d/2)n + M}$$

Replacing n in $f(n)$ by a variable x that can take on all positive values, we obtain a differentiable function that agrees with $f(n)$ when x is integral. Since this function has the same form as $f(n)$, we denote it by $f(x)$. We now find the value of x making $f(x)$ maximum. If there is only one critical value and this value is an integer, it must also be the value of n making $f(n)$ maximum. If it is not an integer, but lies between the consecutive integers n_1 and $n_1 + 1$, one or both of these integers must maximize $f(n)$. Hence we test them to see which makes $f(n)$ greater.

Applying the usual methods we find that $f(x)$ is maximum if $x = \sqrt{2M/d}$. Inserting the numerical values of M and d , we have $\sqrt{2M/d} = \sqrt{66.7} = 8.2$, approximately. The integral value of n that makes $f(n)$ maximum is therefore 8 or 9. Substituting each in the expression for I , we find that 8 yields the greater value.

27. A printer contracts to print 2000 circulars at \$1.00 per hundred. If the number of circulars exceeds 2000, he will deduct 2 cents per hundred on the whole contract, for each hundred printed in excess of 2000. For what number of circulars would the receipts be greatest, and how much would these receipts amount to?

28. A manufacturer who has a monopoly on a certain article produces x units at a

total cost of $ax^2 + bx + c$ dollars. The selling price p is related to the supply by the equation $p = \alpha - \beta x$. Find the output and the price for the maximum profit, if a, b, c, α , and β are positive constants.

29. Solve Exercise 28 if $p = \alpha - \beta x^2$. (See Exercise 10, Art. 3.)

30. A steel plant is capable of producing x tons per day of low-grade steel and y tons per day of high-grade steel, where $y = \frac{a - bx}{c - x}$ in which a, b , and c are positive numbers and $c > x$. If the fixed price of low-grade steel is half that of high-grade steel, find the number of tons of low-grade steel that should be produced per day for maximum profit.

5. Alternative Method. The method described in Art. 4 can be varied to advantage in some cases, as shown in the following example.

EXAMPLE. Find the ratio of the height of a cylindrical can to its radius for the greatest capacity with a given total area.

SOLUTION. As before, we find a function to represent the volume that is to be maximized. This is

$$V = \pi r^2 h. \quad (5.1)$$

Since the total area is a given constant, we write

$$2\pi r^2 + 2\pi r h = A = \text{constant}. \quad (5.2)$$

In this expression, $2\pi r^2 =$ area of the two ends, and $2\pi r h =$ area of the curved surface. By the method of Art. 4 we would use Eq. 5.2 to express h in terms of r , and then insert this in (5.1) to express V as a function of one variable r . Instead, we leave this elimination until later. We differentiate both Eqs. 5.1 and 5.2, thinking of r as the independent variable and considering h as a function of r determined by the implicit relation (5.2). Thus, from (5.1),

$$D_r V = 2\pi r h + \pi r^2 D_r h. \quad (5.3)$$

Differentiating the implicit equation (5.2) connecting r and h ,

$$4\pi r + 2\pi h + 2\pi r D_r h = D_r A = 0. \quad (5.4)$$

We now set the derivative of V equal to 0, assuming that V is to have its maximum value,

$$2\pi r h + \pi r^2 D_r h = 0. \quad (5.5)$$

Equations 5.2, 5.4, and 5.5 form a system of three equations in the three unknowns r, h , and $D_r h$. This system can be solved for any or all of these unknowns, but the process is usually easier than it sounds. Thus, let us eliminate $D_r h$ from (5.4) and (5.5). We multiply (5.4) by r and (5.5) by 2, getting

$$4\pi r^2 + 2\pi r h + 2\pi r^2 D_r h = 0, \quad (5.6)$$

$$4\pi r h + 2\pi r^2 D_r h = 0. \quad (5.7)$$

Subtracting (5.7) from (5.6), we have

$$4\pi r^2 - 2\pi rh = 0. \quad (5.8)$$

Dividing by $2\pi r$, which cannot be 0 if the volume is to be maximum, we obtain

$$2r - h = 0, \text{ or } h/r = 2. \quad (5.9)$$

This method is usually better than that of Art. 4 for the problem of finding the ratio of one variable to another that results in a maximum or minimum. Finding such a ratio is sometimes a good preliminary step toward obtaining the actual values that produce a maximum or minimum. Thus, putting $h = 2r$, from (5.9), into (5.2), we get

$$2\pi r^2 + 4\pi r^2 = A,$$

and

$$r^2 = A/6\pi, \quad r = \sqrt{A/6\pi}.$$

Then, since $h = 2r$,

$$h = 2\sqrt{A/6\pi}.$$

If $A = 40$ sq in., for example,

$$r = \sqrt{40/6\pi} = 1.45 \text{ in., approximately,}$$

and

$$h = 2\sqrt{40/6\pi} = 2.91 \text{ in., approximately.}$$

Whether to use the method just illustrated or that of Art. 4 is a question that cannot be answered by a categorical rule. To a large extent, experience must be the guide. The student should learn both methods in order to have as many working tools as possible. It is worth while to solve some problems by both methods, both to clarify the steps in the method of this article and to have a basis for comparison of the two.

EXERCISES

Solve the following problems by the method illustrated in the above example.

1. Find the dimensions of the rectangle of smallest perimeter whose area is 900 sq ft.

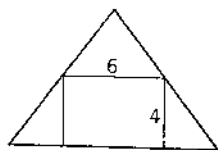


FIG. 14

2. Of all isosceles triangles circumscribed about a rectangle with sides 6 and 4 units, respectively, find the ratio of the altitude to the base in order that the triangle will have the least area, assuming that the longer side of the rectangle lies on the base of the triangle (Fig. 14).

3. A cylinder having as large a volume as possible is inscribed in a cone, as in Fig. 15. If the radius of the base of the cone is 8 and the altitude of the cone is 16, find the ratio of the radius r of the cylinder to its altitude h . Find the dimensions of this cylinder.

4. Find the ratio of r to h in Exercise 3 if, instead of the volume of the cylinder being a maximum, its lateral area is a maximum.

5. What should be the shape of a conical tent (ratio of r to h) in order that it may enclose a given volume with the least amount of canvas?

6. Solve Exercise 10, Art. 4, by the method of this article.

7. Solve Exercise 11, Art. 4, by the method of this article.

8. Solve Exercise 17, Art. 4, by the method of this article.

9. Solve Exercise 20, Art. 4, by the method of this article.

10. Solve Exercise 21, Art. 4, by the method of this article.

11. The cost of the armature of a dynamo is x and the cost of the rest of the dynamo is y . The cost of running the dynamo is c , where $c = 10x + 3y$. The power generated is proportional to xy . If c is fixed find the ratio of x to y to obtain the maximum power.

12. The cost of lining a cylindrical tank with open top is $\frac{3}{2}$ times as much per square foot for the bottom as for the side. If the volume is fixed, find the ratio of the depth to the radius for a minimum cost of lining.

13. A rectangular tank to hold 432 cu ft is to be built with a square base and open top and is to be lined with metal. If it costs 20 cents per sq ft for the lining of the sides and 10 cents per sq ft for the bottom, what should be the dimensions for the least cost for the lining?

14. A rectangular reservoir with a square base and open top, with a capacity of 1125 cu ft, is to be constructed of poured concrete. If the walls are to be 9 in. thick and the base 6 in. thick, what should be the inside dimensions for the least amount of concrete?

15. Find the ratio of the sides of the largest rectangle that can be inscribed in the ellipse $x^2/a^2 + y^2/b^2 = 1$.

16. Find the dimensions of the right circular cylinder of greatest lateral area that can be inscribed in a sphere 6 in. in diameter.

17. The stiffness of a rectangular beam of given length and material varies directly as the product of the breadth and the cube of the depth. Find the shape (ratio of the dimensions) of the stiffest beam that can be sawed from a log whose cross section is a circle.

18. A silo is to be made in the form of an open-topped cylinder surmounted by a conical roof, the inclination of the roof being 45° . The silo, when filled to the top of the cylindrical wall, is to hold 5000 cu ft. If the cost for the base and roof is twice as much per square foot as for the side wall, find the most economical shape and the corresponding dimensions.

19. If, in Exercise 18, the form is to be a cylinder surmounted by a hemispherical roof, and the cost per square foot of erecting the base and side walls is the same,

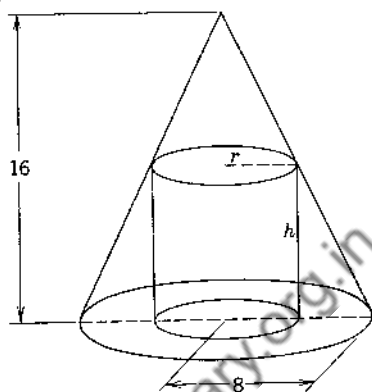


FIG. 15

whereas the roof costs $\frac{5}{2}$ times as much per square foot, find the most economical dimensions.

20. In Fig. 16, find the ratio of AB to AC in order that the triangle ABC shall have the least area, if BC must pass through the fixed point P .

21. The whole area of a handbill is to be 24 sq in. If the margins at the top and bottom are one inch wide and those at the sides $\frac{1}{2}$ in. wide, what should be the height and width in order that the printed matter may occupy the greatest area?

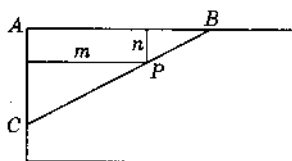


FIG. 16

22. Find the dimensions of the isosceles triangle of least area that circumscribes a circle of radius 3 in.

23. Find the dimensions of the right circular cone of least volume that circumscribes a sphere of radius 6 in.

24. Find the shape of the right circular cone of greatest volume that can be inscribed in a given sphere.

6. Other Methods. The methods by which maximum and minimum problems were solved in this chapter are powerful, and can be employed advantageously for the solution of a great variety of problems. We shall review these methods at several points later on, where we shall not be limited to algebraic functions. One must not hastily conclude, however, that all maximum and minimum problems are to be dealt with in this way. Indeed, the answer can sometimes be found by inspection, whereas an uncritical use of the derived function may lead to an absurd result. For example, let $y = \sqrt{x^2 - 1}$, for $|x| \geq 1$. Clearly, y is least for $|x| = 1$. But if we set $D_x y = 0$, we obtain $x/(\sqrt{x^2 - 1}) = 0$, whence $x = 0$ and y is imaginary. The actual minimum occurs at the end of the range of definition of the function, and there is no minimum in the interior of this range.

Maximum and minimum problems have been the subject of very intensive study for centuries, and have inspired many profound developments in mathematics. Some of the methods devised for special classes of problems depend on the simplest ideas of algebra and geometry. Their study is a profitable supplement to the study of maxima and minima in calculus.*

* See, for example, Courant and Robbins, *What is Mathematics?* Chapter VII.

EXERCISES

1. Show by geometry that the triangle of greatest area, in which two sides are given, is a right triangle.
2. Find the minimum distance from the point $(-4, 0)$ and the parabola $x = y^2$, and show that the result cannot be obtained by the method of Art. 4.
3. A man at A (Fig. 17) wants to get a pail of water from the straight river R and carry it to B in the shortest possible time. Show by geometry that he should follow the route ACB so that $\angle \alpha = \angle \beta$.

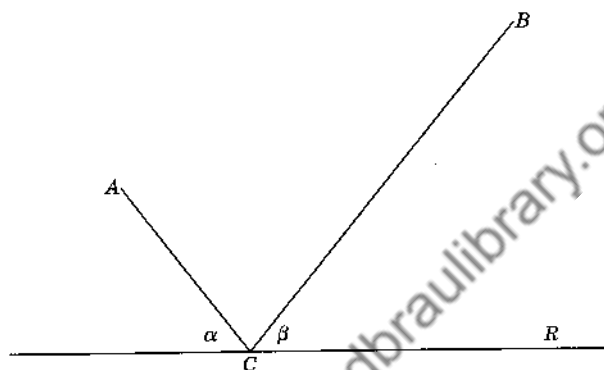


FIG. 17

4. An ellipse is often defined as the locus of points such that the sum of the distances from two fixed points, or foci, is constant. Show that the lines joining the foci to a point of the ellipse meet the ellipse at equal angles. This is the reflection property of the ellipse, whereby light emanating from one focus is always reflected to the other focus. [Hint: Use the result of Exercise 3. The point C is the point of R for which $AC + CB$ is minimum. Now, imagine a line L meeting the ellipse at the point that is joined to the foci. The lines to the foci make equal angles with L if, and only if, L is tangent to the ellipse. For, if L enters the ellipse, the sum of the distances from the foci to a point of L is minimum at an interior point.]

6

The Inverse Rate Problem

1. Introduction. The rate problem raises the following question: if we are given a functional relation between two variables, $y = f(x)$, what is the rate of change of y per unit change of x ? This problem was solved by applying the operation of differentiation to the functional relation, $y = f(x)$. It is natural to expect that it is often necessary to solve the opposite problem, namely, the problem of finding the functional relation between y and x if we know the rate of change of y per unit change of x . For example, we may know the velocity at which a body moves at any time, and wish to express the distance moved as a function of the time. This has already been done in simple cases, in Chapter 4. Thus, suppose v is constantly 20. It is easy to see that $s = 20t$ is a functional relation that satisfies this condition. Also, as observed in Chapter 4, the velocity function $v = 32t + v_0$ might result from the distance equation

$$s = 16t^2 + v_0t + s_0.$$

Similarly, in Chapter 4 we solved the problem of finding the equation of a curve from a formula for its slope, in simple examples. Since the slope problem is the prototype of all rate problems, this problem of finding the equation of the curve from its slope formula is a prototype of the opposite, or *inverse rate problem*. Inverse rate problems have great scientific importance as well as mathematical interest. Often a rate can be determined directly by a laboratory measurement. At other times, scientific reasoning leads to conclusions concerning rates. It is desirable for mathematics to furnish a means of finding the functional relations corresponding to these rates.

Since the rate problem is solved by differentiation, what is now needed is an operation that leads in the opposite direction. That is, we need rules by which we can express y as a function of x when we know the function that is equal to $D_x y$. More briefly, we must learn how to find $f(x)$ from $f'(x)$.

2. Operations and their Inverses. The operation for which we now have need may be called the *inverse* of differentiation. Inverse operations

are by no means unfamiliar. Addition has its inverse, subtraction. Division is the inverse of multiplication. Involution has its inverse, evolution, or root extraction. In each of these pairs of operations, the second is the inverse of the first and the first is the inverse of the second.

The operation of finding a logarithm and that of finding an antilogarithm are learned together, although we conceive the antilogarithm first. Thus, we define the logarithm of N to the base 10 to be the number L that makes the equation, $10^L = N$, true. Here, N is the antilogarithm of L and L is the logarithm of N . In the use of a table of logarithms, looking up the logarithm of a number and looking up the antilogarithm of a given logarithm are inverse operations.

We considered a generalization of inverse operations in Chapter 3, Art. 9. There we thought of y as being any function of x , $y = f(x)$; that is, a value of y could be obtained from a value of x by any of a variety of operations. The inverse function, $x = g(y)$, then, gives directions for reversing the process and finding a value of x from a value of y .

3. Inverse Differentiation. The operation that is the inverse of differentiation we shall call, for the present, *inverse differentiation*, *antidifferentiation*, or finding the *inverse derivative* or *antiderivative*. Rules for this operation are obtained by reading the formulas for differentiation in reverse order. Thus, instead of saying that the derived function of x^2 is $2x$, we may say that a function whose derived function is $2x$ is x^2 . A function from which a certain derived function arises by differentiation is called a *primitive function* of that derived function. Hence, x^2 is a primitive function of $2x$.

Some inverse operations are one valued; e.g., the number whose common logarithm is 2 is 100 and cannot be otherwise. Others lead to ambiguous results; e.g., the angle whose sine is $\frac{1}{2}$ is 30° , 150° , 390° , or any one of various other angles. Inverse differentiation is an operation of the second type. Since $Df(x) = 2x$ if $f(x) = x^2$, $x^2 + 1$, $x^2 + 2$, or $x^2 + \text{any constant}$, we must conclude that any one of these functions is a primitive function of $2x$. It is easily verified that if the differentiation of some function $f(x)$ yields the derived function $g(x)$, then the differentiation of $f(x) + c$, where c is any constant, yields the same derived function. That is, each of the functions obtained from $f(x) + c$ by the assignment of different numbers to c is a primitive function of $g(x)$. The fact that by the device of adding an arbitrary constant to one primitive function of a given function, we include all possible primitive functions, is expressed by the following theorem, the proof of which we postpone to Chapter 12.

THEOREM. If two functions have the same derived function, they differ at most by an additive constant.

For conciseness, a symbol for the inverse derivative of a function is

needed, comparable to the symbol for the derivative. We take D^{-1} for this symbol, and write

$$D^{-1}(2x) = x^2 + c.$$

Care should be taken not to confuse this use of the superscript -1 with the exponent -1 . We read the expression $D^{-1}(2x)$: "inverse D $2x$," "inverse derivative of $2x$," "antiderivative of $2x$," or "primitive function of $2x$."

It should be clear that $DD^{-1}u(x) = u(x)$, since the derived function of a function whose derived function is $u(x)$ is surely $u(x)$. It is not true that $D^{-1}Du(x) = u(x)$, however, because of the ambiguity of the inverse differentiation process. Rather, we must write

$$D^{-1}Du(x) = u(x) + c.$$

Each statement connecting a function with its derived function can be written in two ways. Thus, for example, we can write

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

and

$$D(x^n + c) = nx^{n-1} \quad \text{or} \quad D^{-1}(nx^{n-1}) = x^n + c.$$

Similarly, we could write an inverse derivative theorem corresponding to each of the theorems listed in Chapter 3, Art. 14. We call particular attention, however, to the following.

$$D^{-1}(0) = c, \tag{3.1}$$

$$D^{-1}(u + v) = D^{-1}u + D^{-1}v + c, \tag{3.2}$$

and

$$D^{-1}(au) = aD^{-1}u + c, \tag{3.3}$$

where a is a constant. Each of these statements is justified by direct differentiation. Thus,

$$D[D^{-1}(0)] = 0 \quad \text{and} \quad Dc = 0;$$

$$D[D^{-1}(u + v)] = u + v,$$

and

$$D[D^{-1}u + D^{-1}v + c] = D[D^{-1}u] + D[D^{-1}v] + Dc = u + v;$$

and

$$D[D^{-1}(au)] = au,$$

$$D(aD^{-1}u + c) = aDD^{-1}u + Dc = au.$$

In the above equations D symbolizes differentiation with respect to x , and u and v are functions of x . Similarly, D^{-1} represents inverse differentiation with respect to x , i.e., finding a function whose derivative with

respect to x is given. When it is not clear what variable is independent, its symbol can be attached to D^{-1} to make this explicit. Thus, we might write

$$D_x^{-1}(au) = aD_x^{-1}u + c.$$

An inverse derivative formula corresponding to formula $D-1$ can be written in the form

$$D^{-1}(nx^{n-1}) = x^n + c, \quad (3.4)$$

provided that $n \neq 0$. This can be put into more useful form as follows. First, we have

$$D^{-1}[(n+1)x^n] = x^{n+1} + c', \quad \text{if } n \neq -1. \quad (3.5)$$

This is justifiable, since (3.4) is correct for any rational power, except 0, in particular for $n+1$. The constant is denoted by c' , reserving c for the final formula. Now, by (3.3),

$$D^{-1}[(n+1)x^n] = (n+1)D^{-1}x^n.$$

Hence, (3.5) becomes

$$(n+1)D^{-1}x^n = x^{n+1} + c'.$$

Since $n+1 \neq 0$, we may divide by it, and obtain

$$D^{-1}x^n = \frac{x^{n+1}}{n+1} + c, \quad (I-1)$$

where $c = c'/(n+1)$. This is the fundamental power formula for inverse differentiation, corresponding to formula $D-1$ of Chapter 3. Note particularly that this formula holds only if $n+1 \neq 0$, i.e., $n \neq -1$.

We have seen that the result of any inverse differentiation must include an arbitrary additive constant. The form of this constant is immaterial as long as it is free to assume any real value. Thus, as an example of formula 3.3, we write

$$D^{-1}(6x^4) = 6D^{-1}x^4 + c.$$

Using (I-1),

$$D^{-1}x^4 = x^5/5 + c',$$

where c' is independent of c . Combining these results,

$$D^{-1}(6x^4) = 6(x^5/5 + c') + c = 6x^5/5 + 6c' + c.$$

The constant here has the form $6c' + c$ rather than a single letter, but this makes no essential difference. The important point is that we need a quantity added that can take on any constant value, and $6c' + c$ and c will serve the purpose equally well. Hence, we write the result simply

$6x^5/5 + c$. The presence of the arbitrary constant makes our results indefinite and ambiguous. This is an advantage rather than a handicap. Because of it the results can be adapted to a variety of needs that could not otherwise be met. This will be made clear in the solution of inverse rate problems in Arts. 5-7.

Equations 3.2 and 3.3, together with formula I-1, will make it possible to find the inverse derivatives of all polynomials and of many fractional and irrational expressions.

EXAMPLE 1. Given $Dy = 3x^2$, find y .

SOLUTION. Applying (3.3),

$$y = 3D^{-1}x^2.$$

Applying formula I-1,

$$\begin{aligned} y &= 3(x^3/3 + c) \\ &= x^3 + 3c. \end{aligned}$$

Since $3c$ can stand for any constant as well as and no better than c , we can just as well write

$$y = x^3 + c.$$

EXAMPLE 2. Given $Dy = 3x^2 + 2x - 1$, find y .

SOLUTION. Applying (3.2), (3.3), and (I-1),

$$\begin{aligned} y &= D^{-1}(3x^2) + D^{-1}(2x) - D^{-1}(1) \\ &= x^3 + c_1 + x^2 + c_2 - x - c_3. \end{aligned}$$

Since $c_1 + c_2 - c_3$ is always a constant, we may put

$$c_1 + c_2 - c_3 = c$$

and write

$$y = x^3 + x^2 - x + c.$$

EXAMPLE 3. Given $Dy = 1/x^3$, find y .

SOLUTION. Writing $1/x^3 = x^{-3}$, and applying formula I-1, we have

$$y = D^{-1}(x^{-3}) = \frac{x^{-2}}{-2} + c = -\frac{1}{2x^2} + c.$$

EXAMPLE 4. Given $Dy = \sqrt{x} - \frac{1}{\sqrt{x}}$, find y .

SOLUTION. Writing $\sqrt{x} = x^{\frac{1}{2}}$, we have

$$Dy = x^{\frac{1}{2}} - x^{-\frac{1}{2}}.$$

Now, applying (3.2) and (I-1),

$$\begin{aligned} y &= D^{-1}(x^{\frac{3}{2}} - x^{-\frac{1}{2}}) = \frac{x^{\frac{3}{2}}}{\frac{3}{2}} + c_1 - \frac{x^{\frac{1}{2}}}{\frac{1}{2}} - c_2 \\ &= \frac{2}{3} \cdot x^{\frac{3}{2}} - 2x^{\frac{1}{2}} + c, \quad \text{where } c = c_1 - c_2 \\ &= \frac{2}{3} \cdot x\sqrt{x} - 2\sqrt{x} + c. \end{aligned}$$

EXERCISES

Obtain y as a function of x in each of Exercises 1-25.

1. $Dy = 3x.$
2. $Dy = 5x^4.$
3. $Dy = x^7.$
4. $Dy = 2x + 3.$
5. $Dy = x^2 - 3x - 5.$
6. $Dy = x^4 - x^2 + x^2.$
7. $Dy = \frac{3}{2}x^2 - \sqrt{2x} + \frac{1}{2}.$
8. $Dy = 5x^3 - \frac{1}{3}x^2 + \frac{2}{3}x - 11.$
9. $Dy = 10x^7 + 14x^6 - 18x^5 - 15x^4 + 6x^3 - 21x^2 + 8x + 40.$
10. $Dy = (x^3 - 1)^2.$
11. $Dy = 1/x^3.$
12. $Dy = x - x^{-2}.$
13. $Dy = \sqrt{x} + 1/\sqrt{x}.$
14. $Dy = 5x^4 + 3/x^4.$
15. $Dy = 2x^3 - 3x^{-3} + 4 + 3x - 4x^2.$
16. $Dy = 2x^{\frac{1}{2}}.$
17. $Dy = 3x^{\frac{1}{2}} - 2x^{\frac{1}{2}} + 1.$
18. $Dy = 10x^{\frac{3}{2}} + 3x^{-\frac{1}{2}}.$
19. $Dy = 9\sqrt{x} + 4x - 5x\sqrt{x}.$
20. $Dy = 5x^4 + 4x^{\frac{1}{2}} + x.$
21. $Dy = 2x^{-\frac{1}{2}} - 4x^{-\frac{1}{2}} + 3 + 4x^{\frac{1}{2}} - 5x^{\frac{3}{2}}.$
22. $Dy = 3\sqrt{x} - 8\sqrt[3]{x}.$
23. $Dy = 1/\sqrt[3]{x} + 2/\sqrt{x} - 1/\sqrt{x}.$
24. $Dy = \sqrt[3]{x^2} + 1/x^3 + \sqrt{x^3}.$
25. $Dy = 1/\sqrt[3]{x} + 1/\sqrt{x} + \frac{1}{2} + x + \sqrt{x} + \sqrt[3]{x^2}.$

Show that each of the statements in Exercises 26-30 is correct by carrying out the expressed inverse differentiation on the left-hand side.

26. $D^{-1}(x+1)^2 = \frac{1}{3}(x+1)^3 + c.$ [Hint: write $(x+1)^2 = x^2 + 2x + 1$.]
27. $D^{-1}(x-1)(x+1) = \frac{1}{3}x(x^2-3) + c.$
28. $D^{-1}\left(\frac{1}{x^n}\right) = -\frac{1}{(n-1)x^{n-1}} + c, \text{ if } n \neq 1.$
29. $D^{-1}(x^{n+1}) = \frac{nx^{(n+1)/n}}{n+1} + c, \text{ if } n \neq 0 \text{ or } -1.$
30. $D^{-1}(x+1)^n = \frac{(x+1)^{n+1}}{n+1} + c, \text{ if } n = \text{a positive integer.}$ [Hint: Expand $(x+1)^n$

by the binomial theorem, and express the arbitrary constant in the form $1/(n+1) + c$.]

Verify Exercises 31-40 by differentiating the right-hand side.

31. $D^{-1}8x(x^2-1)^3 = (x^2-1)^4 + c.$
32. $D^{-1}4(9x^2+1)(3x^3+x)^3 = (3x^3+x)^4 + c.$
33. $D^{-1}\frac{x}{\sqrt{x^2+1}} = \sqrt{x^2+1} + c.$

$$34. D^{-1} \frac{3x^2}{2\sqrt{x^3+5}} = \sqrt{x^3+5} + c.$$

$$35. D^{-1}[6x(x^2+1)^2(2x^3+1)(7x^3+5x+1)] = (x^2+1)^3(2x^3+1)^2 + c.$$

$$36. D^{-1} \frac{x^2+2x}{(x+1)^2} = \frac{x^2}{x+1} + c.$$

$$37. D^{-1} \frac{5x+2}{6\sqrt{x+1}\sqrt[3]{x^2}} = \sqrt{x+1}\sqrt[3]{x} + c.$$

$$38. D^{-1} \frac{2x}{\sqrt{x^2+1}\sqrt{(x^2-1)^3}} = -\frac{\sqrt{x^2+1}}{\sqrt{x^2-1}} + c.$$

$$39. D^{-1} \frac{x+2}{6(1-x)\sqrt[3]{x^2}\sqrt{x-1}} = \frac{\sqrt[3]{x}}{\sqrt{x-1}} + c.$$

40. $D^{-1} \frac{1}{x^2} = \frac{x-1}{x} + c$. Explain the apparent discrepancy between this result and that obtained as follows.

$$D^{-1} \left(\frac{1}{x^2} \right) = D^{-1}(x^{-2}) = \frac{x^{-1}}{-1} + c = -\frac{1}{x} + c.$$

4. The General Power Formula. The chain rule for differentiation enabled us to write the derived function of any power of a function of x . Thus, we have formula $D-1a$,

$$D_x[u(x)]^n = nu^{n-1}D_xu. \quad (D-1a)$$

From this we can derive a general power formula for inverse differentiation. From $(D-1a)$ we conclude that

$$D_x u^{n+1} = (n+1)u^n D_x u,$$

and

$$D_x \frac{u^{n+1}}{n+1} = u^n D_x u, \quad \text{if } n \neq -1.$$

From the theorem in Art. 3, the primitive functions of these equal functions must be the same except, at most, for an additive constant. That is,

$$D_x^{-1} D_x \frac{u^{n+1}}{n+1} = D_x^{-1}(u^n D_x u) + c', \quad \text{if } n \neq -1,$$

or, writing in reverse order and applying the inverse operations $D_x^{-1} D_x$,

$$D_x^{-1}(u^n D_x u) = \frac{u^{n+1}}{n+1} + c, \quad \text{if } n \neq -1, \quad (I-1a)$$

where $c = -c'$. Formula $I-1a$ is a generalization of $(I-1)$.

The chain rule for differentiation tells us that we can differentiate a power of a function of x , viz., u^n , by the same rule that is used for x^n , if we then multiply the result by $D_x u$. Reversing the procedure, formula I-1a tells us that we can write the inverse derivative of a power of a function of x , viz., u^n , by the same rule that is used for x^n , if it is accompanied by the factor $D_x u$. The necessity that the factor $D_x u$ be present must not be overlooked.

Some examples will help to show the utility of formula I-1a and also the care that must be exercised in its use.

EXAMPLE 1. Evaluate $D_x^{-1}(x^2 + 1)^5 2x$.

SOLUTION. We can put $x^2 + 1 = u$ and $2x = D_x u$. We then have

$$D_x^{-1}(x^2 + 1)^5 2x = D_x^{-1}[u^5 D_x u] = \frac{u^6}{6} + c = \frac{(x^2 + 1)^6}{6} + c.$$

The notations, D_x and D_x^{-1} , using the subscript x , are necessary only when an intermediate variable such as u is actually expressed. In the above example we might write

$$D^{-1}(x^2 + 1)^5 2x = \frac{(x^2 + 1)^6}{6} + c.$$

Thus, we conceive $x^2 + 1$ as being a new variable and $2x$ as its derivative, without assigning a letter to this variable. Although we shall find that it is not always necessary to rewrite our expressions inserting u and $D_x u$, in order to avoid mistakes it is often best to do so.

EXAMPLE 2. Evaluate $D^{-1}(3x^4 - 2)^7 12x^3$.

SOLUTION. Putting $3x^4 - 2 = u$, we have $D_x u = 12x^3$. Hence, we have to evaluate $D_x^{-1}[u^7 D_x u]$. By formula (I-1a) the result is $u^8/8 + c$. Finally, replacing u by its value $3x^4 - 2$, we have

$$D^{-1}(3x^4 - 2)^7 12x^3 = \frac{(3x^4 - 2)^8}{8} + c.$$

EXAMPLE 3. Evaluate $D^{-1}(2x^5 + 3x)^{10}(10x^4 + 3)$.

SOLUTION. Putting $2x^5 + 3x = u$, we have $D_x u = 10x^4 + 3$. From formula I-1a we have

$$\begin{aligned} D^{-1}(2x^5 + 3x)^{10}(10x^4 + 3) &= D_x^{-1}[u^{10} D_x u] = \frac{u^{11}}{11} + c \\ &= \frac{(2x^5 + 3x)^{11}}{11} + c. \end{aligned}$$

The fact that formula I-1a can be used only when, in addition to the power of a quantity, we have the derivative of that quantity as a factor, suggests that the use of the formula is very restricted. This restriction is

not as great as may seem to be the case at first, since, if some factor other than $D_x u$ is present instead of $D_x u$, we can sometimes change this factor to $D_x u$ and compensate for the change by a simple means which does not interfere with the operation of inverse differentiation.

EXAMPLE 4. Evaluate $D^{-1}(x^3 + 3)^5 x^2$.

SOLUTION. If we put $x^3 + 3 = u$, we have $D_x u = 3x^2$. Thus, we have to evaluate $D_x^{-1} \left[u^5 \frac{D_x u}{3} \right]$. This can be done, using formula 3.3. Thus

$$\begin{aligned} D_x^{-1} \left[u^5 \frac{D_x u}{3} \right] &= \frac{1}{3} D_x^{-1} [u^5 D_x u] \\ &= \frac{1}{3} \frac{u^6}{6} + c. \end{aligned}$$

The answer to our problem is therefore

$$\frac{1}{3} \frac{(x^3 + 3)^6}{6} + c, \quad \text{or} \quad \frac{1}{18} (x^3 + 3)^6 + c.$$

EXAMPLE 5. Evaluate $D^{-1}(3x^2 - 5)^3 5x$.

SOLUTION. Putting $3x^2 - 5 = u$, $D_x u = 6x$, we have

$$\begin{aligned} D^{-1}(3x^2 - 5)^3 5x &= D_x^{-1} [u^3 \frac{5}{6} D_x u] \\ &= \frac{5}{6} D_x^{-1} [u^3 D_x u] \\ &= \frac{5}{6} \frac{u^4}{4} + c \\ &= \frac{5}{24} u^4 + c = \frac{5}{24} (3x^2 - 5)^4 + c. \end{aligned}$$

EXAMPLE 6. Evaluate $D^{-1}x\sqrt{x^2 + 4}$.

SOLUTION. Putting $x^2 + 4 = u$, $D_x u = 2x$, we have

$$\begin{aligned} D^{-1}x\sqrt{x^2 + 4} &= D_x^{-1} [u^{\frac{1}{2}} \frac{1}{2} D_x u] \\ &= \frac{1}{2} D_x^{-1} [u^{\frac{1}{2}} D_x u] \\ &= \frac{1}{2} \frac{u^{\frac{3}{2}}}{\frac{3}{2}} + c \\ &= \frac{1}{3} u^{\frac{3}{2}} + c = \frac{1}{3} (x^2 + 4)^{\frac{3}{2}} + c. \end{aligned}$$

Each of the inverse derivatives in Examples 1-5 could have been evaluated by the alternative process of expanding the power, multiplying

in the other factor, and finally applying methods of Art. 3. Thus, in Example 5,

$$\begin{aligned}(3x^2 - 5)^3 5x &= (27x^6 - 135x^4 + 225x^2 - 125)5x \\ &= 135x^7 - 675x^5 + 1125x^3 - 625x.\end{aligned}$$

Applying inverse differentiation to this expression, we get

$$\frac{135x^8}{8} - \frac{675x^6}{6} + \frac{1125x^4}{4} - \frac{625x^2}{2} + c.$$

The student may show that the result is equivalent to that obtained above.

The alternative of expanding the power of the binomial is not available in Example 6, nor in Examples 7 and 8, which follow.

EXAMPLE 7. Evaluate $D^{-1} \frac{3x^4}{(2x^5 - 7)^2}$.

SOLUTION. Putting $2x^5 - 7 = u$, $D_x u = 10x^4$, we have

$$\begin{aligned}D^{-1} \frac{3x^4}{(2x^5 - 7)^2} &= D_x^{-1} [u^{-2} \frac{3}{10} D_x u] \\ &= \frac{3}{10} D_x^{-1} [u^{-2} D_x u] \\ &= \frac{3}{10} \frac{u^{-1}}{-1} + c \\ &= -\frac{3}{10} \frac{1}{2x^5 - 7} + c.\end{aligned}$$

EXAMPLE 8. Evaluate $D \frac{6x - 9}{\sqrt{x^2 - 3x + 2}}$.

SOLUTION. Putting $x^2 - 3x + 2 = u$, $D_x u = 2x - 3$, we have

$$\begin{aligned}D^{-1} \frac{6x - 9}{\sqrt{x^2 - 3x + 2}} &= D_x^{-1} [u^{-\frac{1}{2}} 3 D_x u] \\ &= 3 D_x^{-1} [u^{-\frac{1}{2}} D_x u] \\ &= 3 \frac{u^{\frac{1}{2}}}{\frac{1}{2}} + c \\ &= 6\sqrt{x^2 - 3x + 2} + c.\end{aligned}$$

It should be observed that a factor other than $D_x u$ can be changed to $D_x u$, so that formula I-1a can be used, only when this change is brought about by multiplication by a constant. Example 9 illustrates a futile attempt to make the change in a different situation.

EXAMPLE 9. Evaluate $D^{-1}(x^3 + 1)^{\frac{1}{3}}x$.

SOLUTION. If we set $x^3 + 1 = u$, we have $D_x u = 3x^2$. We can now write

$$\begin{aligned} D^{-1}(x^3 + 1)^{\frac{1}{3}}x &= D^{-1} \left[(x^3 + 1)^{\frac{1}{3}} \frac{1}{3x} 3x^2 \right] \\ &= D_x^{-1} \left[\frac{1}{3x} u^{\frac{1}{3}} D_x u \right]. \end{aligned}$$

At this stage we cannot proceed further, for the factor $\frac{1}{3x}$ cannot be moved in front of the symbol D_x^{-1} , since formula 3.3 allows only a constant to be moved in this way.

EXERCISES

Evaluate the inverse derivatives expressed in Exercises 1-6.

1. $D^{-1}2x\sqrt{x^2 + 1}$.

2. $D^{-1}3x\sqrt{x^2 - 1}$.

3. $D^{-1}5x^3(3x^4 - 2)^{\frac{2}{3}}$.

4. $D^{-1} \frac{3}{\sqrt[3]{3x + 5}}$.

5. $D^{-1} \frac{(3 - \sqrt{x})^{\frac{1}{2}}}{\sqrt{x}}$.

6. $D^{-1} \frac{1}{x^2} \left(\frac{1}{x} + 2 \right)^3$.

Obtain y as a function of x in each of Exercises 7-12.

7. $Dy = 2\sqrt{x-1}$.

8. $Dy = (3x^3 + 5)^{13}3x^2$.

9. $Dy = x(1 - x^2)^7$.

10. $Dy = -\frac{\sqrt{1/x - 2}}{x^2}$.

11. $Dy = \frac{x^2}{(4 - x^3)^{\frac{1}{2}}}$.

12. $Dy = (x^n + 1)^n x^{n-1}$, if $n \neq -1$.

13. Show by differentiation that the result obtained by completing Example 9 in the following way is incorrect.

$$\begin{aligned} D^{-1}(x^3 + 1)^{\frac{1}{3}}x &= \frac{1}{3x} D_x^{-1}[u^{\frac{1}{3}} D_x u] = \frac{1}{3x} \frac{u^{\frac{3}{2}}}{\frac{3}{2}} + c \\ &= \frac{2}{9x} u^{\frac{3}{2}} + c \\ &= \frac{2}{9x} (x^3 + 1)^{\frac{3}{2}} + c. \end{aligned}$$

5. Application to Motion in a Straight Line. If a body moves in a straight line, we have seen that its velocity is given by the derivative, $v = D_t s$, where s = the number of units of distance moved in t units of time. Consequently, we can use the inverse derivative to find the relation between s and t when the velocity is known as a function of t .

EXAMPLE 1. Find the equation of motion (relation between s and t) of a body that moves so that the number of units of velocity is equal to the number of units of time.

SOLUTION. We write

$$v = D_t s = t.$$

Therefore,

$$s = D^{-1}t = t^2/2 + c.$$

The equation of motion is still indefinite, since the constant c is undetermined. But if we know s for one value of t we can determine c and, hence, express the equation of motion precisely. Often we know the initial condition, that is, we know s when $t = 0$. Thus, if $s = 0$ when $t = 0$, it follows that

$$0 = 0^2/2 + c, \quad \text{whence} \quad c = 0, \quad \text{and} \quad s = t^2/2.$$

EXAMPLE 2. Find the equation of motion of a body starting from rest, that is, with initial velocity equal to 0, if the acceleration is constant and equal to 32 units.

SOLUTION. Here we have

$$D_t v = 32.$$

Then,

$$v = D^{-1}32 = 32t + c_1.$$

Since the initial velocity is 0, $0 = 32 \cdot 0 + c_1$, $c_1 = 0$, and $v = 32t$.

Now,

$$v = D_t s = 32t.$$

Then,

$$s = D^{-1}32t = 16t^2 + c_2.$$

If s is measured from the point where the body is located when $t = 0$, we have $s = 0$ when $t = 0$, and, consequently, $c_2 = 0$. Therefore, $s = 16t^2$.

The conditions of this example are those of a body falling from a position of rest under the influence of gravity, which causes a constant acceleration of 32 ft per sec per sec if air resistance is neglected. Hence, the equation $s = 16t^2$ expresses the distance in feet that the body falls in t seconds. Compare this example and Example 3, which follows, with the discussion in Chapter 4, Art. 4.

EXAMPLE 3. Find the velocity v and distance s in terms of the time t of a body moving subject to a constant acceleration a , if the initial velocity is v_0 and the initial distance is s_0 .

SOLUTION. We have $a = D_t v = \text{constant}$, therefore,

$$v = D_t^{-1}a = at + c_1.$$

Since $v = v_0$ when $t = 0$,

$$c_1 = v_0 \quad \text{and} \quad v = at + v_0.$$

Now,

$$s = D_t^{-1}v = D_t^{-1}(at + v_0) = \frac{a}{2}t^2 + v_0t + c_2.$$

Since $s = s_0$ when $t = 0$,

$$c_2 = s_0, \quad \text{and} \quad s = \frac{a}{2}t^2 + v_0t + s_0.$$

The acceleration due to gravity at the earth's surface is approximately 32 ft per sec per sec downward. Thus, if we think of the positive direction as upward, $a = -32$, and the above equation becomes

$$s = -16t^2 + v_0t + s_0.$$

This would be the equation that would apply if, for example, a body is thrown directly upward with velocity v_0 ft per sec from a point s_0 feet from the earth's surface, air resistance being neglected.

EXERCISES

1. Find the number of feet a body moves in a straight line in t seconds if the velocity in feet per second is equal to

(a) $20t$.

(b) $100 + 10t$.

(c) $30 + 20t - 15t^2$.

2. If the initial velocity of a body is 0, find the velocity after t seconds if the acceleration in feet per second per second is equal to

(a) 0.

(b) $5 + 4t$.

(c) $t^3 - 6t^2 + 12t - 5$.

3. If $s = 10$ ft and $v = 4$ ft/sec when $t = 0$, find s at the end of t seconds if the acceleration in ft/sec² is equal to

(a) 10.

(b) -6 .

(c) $2t - 4$.

(d) $t^2 + \frac{1}{\sqrt{t+1}}$.

4. The velocity of a body starting from rest is $6t^2$ ft per sec after t seconds. (a) How far from the starting point will it be in 4 sec? (b) What is the acceleration at the end of 2 sec? (c) In how many seconds will it pass over a distance of 432 ft?

5. A body starts from a point 41.6 ft from a given point P with a speed of 2 ft per sec, and moves directly toward P with its acceleration t seconds after starting numerically equal to $.3t$ ft per sec and directed toward P . Find the time required for the body to reach P and its speed at that time.

6. Curvilinear Motion. According to Newton's first law of motion, a body remains at rest or moves with a uniform velocity in a straight line unless acted upon by a force. Such a force produces an acceleration, which is the rate of change of velocity. If the force acts in the direction of the motion, this change is a speeding up or a slowing down of the motion in its straight line path. We considered this case in Art. 5. But

the force may also be directed differently from that of the motion, whereupon it changes the direction of the motion. Thus, in Fig. 1, if a body at O moving in the direction OA is acted upon by a force F in the direction OB , the effect will be to deflect the body toward OB . If, as the body continues to move, this same force is continually applied, the body will be deflected more and more toward OB , that is, its path will be a curve

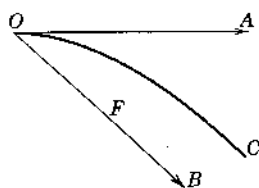


FIG. 1

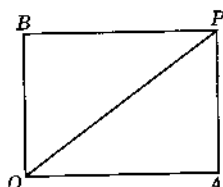


FIG. 2

OC lying between the original direction OA and the direction OB . The study of a situation like this is facilitated by the use of *components*.

If a body moves along the path OP (Fig. 2), it is clear that it moves the horizontal distance OA and the vertical distance OB , simultaneously. We call OA and OB the horizontal component and vertical component, respectively, of this motion. Similarly, in Fig. 3, the arrow labeled v represents the actual velocity of the motion from O to P , which we shall

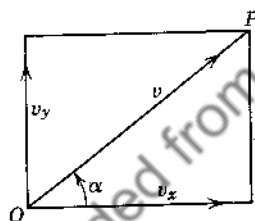


FIG. 3

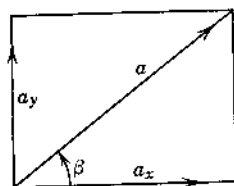


FIG. 4

suppose to be uniform, its length representing the magnitude of the velocity and its direction agreeing with that of the velocity. The arrows labeled v_x and v_y , respectively, represent the simultaneous uniform horizontal velocity and vertical velocity, respectively. Thus, if $v = 5$ (for example, 5 ft per sec), v_x and v_y might be 4 and 3, respectively. Moreover, if v is directed as shown in the figure, v_x must be directed to the right and v_y must be directed upward. The directed quantities v_x and v_y are called the *horizontal* and *vertical components*, respectively, of the directed quantity v .

In like manner, if the motion is accelerated by a constant amount in the direction of the motion, we may represent this acceleration by an arrow a (Fig. 4), the length of a representing the amount of the acceleration,

and the direction of the arrow agreeing with the direction of the acceleration. As with velocity, we imagine the body to be accelerated an amount a_x horizontally and an amount a_y vertically, where a_x and a_y are the perpendicular sides of the right triangle having a as its hypotenuse. The quantity a_x directed horizontally to the right and a_y directed vertically upward are the horizontal and vertical components, respectively, of the acceleration a , directed as shown in the figure.

The directed distances OP , OA , and OB , the directed velocities v , v_x , and v_y , and the directed accelerations a , a_x , and a_y , are examples of *vector* quantities. In fact, we can define a *vector* as a directed quantity. Since a velocity and an acceleration are always directed, they are vector quantities, and it is quite unnecessary to speak of a *directed* velocity or a *directed* acceleration. It is not customary to think of distances as directed unless so specified. We shall use absolute value signs to denote the magnitude of a vector. Thus, we may write $|v|$ when we wish to refer to the amount of v without regard to its direction.

From the definitions of the horizontal and vertical components of the vectors v and a , it should be clear that

$$|v| = \sqrt{v_x^2 + v_y^2}, \quad (6.1)$$

$$|a| = \sqrt{a_x^2 + a_y^2}, \quad (6.2)$$

$$v_x = |v| \cos \alpha, \quad v_y = |v| \sin \alpha, \quad \frac{v_y}{v_x} = \tan \alpha, \quad (6.3)$$

$$a_x = |a| \cos \beta, \quad a_y = |a| \sin \beta, \quad \frac{a_y}{a_x} = \tan \beta. \quad (6.4)$$

In these equations α and β are the angles made with the positive x -axis by v and a , respectively. Although in the present example $\alpha = \beta$, different angles will generally appear in Eqs. 6.3 and 6.4, since the velocity and acceleration are not always directed alike.

As pointed out at the beginning of this discussion, a body moves in a curved path if there is a force acting in a direction different from that of the motion. This means that the path is curved at any point where the acceleration vector and the velocity vector are directed differently. Here it is pertinent to ask just what is meant by the velocity of the body *at a point* of its curved path. We shall obtain an answer to this question in precisely the same way that we formulated the idea of an instantaneous speed in Chapters 1 and 2. Intuitively we know that a body can move in a curve, and we believe there is a definite velocity associated with it at each point of its curved path, varying in amount and direction as it moves along the path. But we must give this idea a clearer formulation than is afforded by intuition.

Instead of thinking of the body as moving along the curve from P to Q_1 (Fig. 5), let us imagine that this motion is replaced by a uniform rectilinear motion along the chord PQ_1 . That is, we consider a substitute motion that would carry the body from P to Q_1 at a uniform speed along the chord in the same time that it actually takes in the true motion.

There is a definite velocity vector v associated with this substitute motion, having components v_x and v_y . Now, think of a different substitute uniform motion that would carry the body along the chord from P to a point Q_2 on the curve and closer to P than Q_1 , in the same time that it actually takes in the true motion to go from P to Q_2 . Again there is a definite velocity vector associated with this second substitute motion. In the same way we may consider a sequence of substitute uniform motions carrying the body from P to $Q_1, Q_2, \dots, Q_n, \dots$, respectively, with Q_n approaching P along the curve. This sequence leads to a sequence of vectors $v^{(1)}, v^{(2)}, \dots, v^{(n)}$ with corresponding sequences of horizontal components $v_x^{(1)}, v_x^{(2)}, \dots, v_x^{(n)}, \dots$ and vertical components $v_y^{(1)}, v_y^{(2)}, \dots, v_y^{(n)}, \dots$ (Fig. 6).

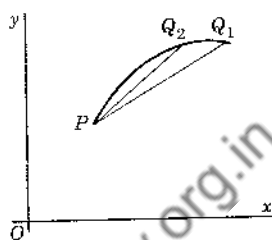


FIG. 5

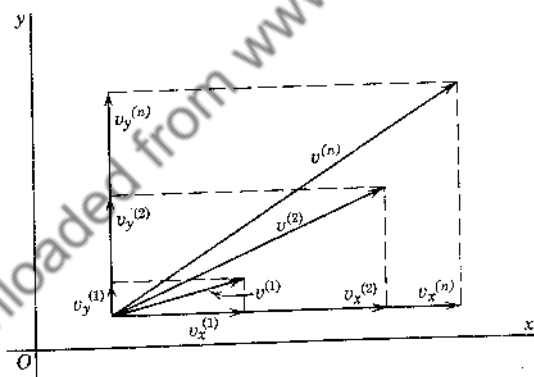


FIG. 6

It should be realized, of course, that the length of the chord, PQ_2 , for example, has no bearing on the magnitude of the substitute velocity vector corresponding to this chord. The chord PQ_2 is simply the straight-line path between two points, whereas the length of the velocity vector represents the rapidity of the motion.

It is to be expected that the sequence of horizontal components $v_x^{(1)}, v_x^{(2)}, \dots, v_x^{(n)}, \dots$ will tend to a limit as n tends to infinity, and that

the sequence of vertical components will also tend to a limit.* These limits, which we denote by v_x and v_y , are taken as the horizontal and vertical components, respectively, of a vector v , where

$$|v| = \sqrt{v_x^2 + v_y^2}, \quad \sin \alpha = \frac{v_y}{|v|}, \quad \cos \alpha = \frac{v_x}{|v|}, \quad \tan \alpha = \frac{v_y}{v_x}.$$

We emphasize that the limiting operations described *define* the two quantities v_x and v_y , which, in turn, define the vector v . As with straight-line motion, these results meet our intuitive requirements of instantaneous velocity in curvilinear motion, and, consequently, we define this velocity as the vector v .

Clearly, the process described is the same as that used to determine a rectilinear velocity, where $v = D_t s$. In the present example, $v_x = D_t X(t)$, where $X(t)$ represents the horizontal distance moved in time t . Similarly, $v_y = D_t Y(t)$, where $Y(t)$ represents the vertical distance moved in time t .

In precisely the same way we define the instantaneous acceleration of a body moving in a curve. Thus, $a_x = D_t v_x$, $a_y = D_t v_y$, and a is the vector acceleration determined by these two components.

7. Distance along a Curve. If we think of s as the distance the body moves along the path from some fixed point of departure, the velocity in the path, v_s , is the same as v , and $|v_s| = \sqrt{v_x^2 + v_y^2}$. We are therefore led to write

$$|D_t s| = \sqrt{(D_t X)^2 + (D_t Y)^2},$$

and

$$D_t s = \pm \sqrt{(D_t X)^2 + (D_t Y)^2}, \quad (7.1)$$

where the sign depends upon the choice of positive direction along the curve.

The quantity on the right-hand side of (7.1) is a function of t that we may call $\phi(t)$. Then

$$s = D^{-1}\phi(t). \quad (7.2)$$

Equation 7.2 defines distance along a curve. The arbitrary constant that is attached to the primitive function is determined by the point of the curve from which the body is considered to start (see Example 3, Art. 8). It is well to be reminded that the elementary definition of length applies only to straight lines, as pointed out in Chapter 1. The problem of defining and calculating the length of a curve is discussed more fully in Chapter 14. At the present time, this problem can be solved completely

* This is true if the path is a smooth curve, i.e., one with no sharp corners or sudden breaks.

only for carefully selected curves. Usually, the inverse derivatives required by Eq. 7.2 cannot be found by methods studied thus far.

8. Examples of Curvilinear Motion. When a body moves in a curved path, its position is usually given by parametric equations, $X = f(t)$, $Y = g(t)$, if the path is a curve in the xy -plane. There is a good reason for this. For we usually find its path from a knowledge of the two components of velocity, v_x and v_y , which are, in turn, found from the two components of acceleration, a_x and a_y . Since a_x and a_y are usually functions of the time t , so also are v_x and v_y and the coordinate distances X and Y of the body, measured from some fixed point.

EXAMPLE 1. A projectile is given an initial velocity of 800 ft per sec inclined at an angle $\pi/6$ with the horizontal. Find the parametric equations of its path, neglecting air resistance. Find the Cartesian equation of the path.

SOLUTION. From Fig. 7 we see that the horizontal and vertical components of the initial velocity are $800 \cos \pi/6 = 693$, and $800 \sin \pi/6$

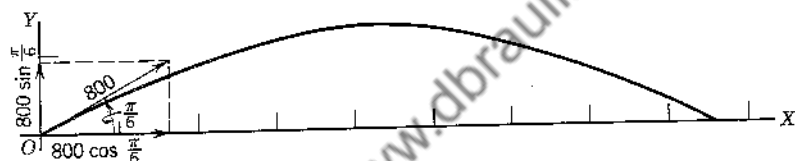


FIG. 7

$= 400$, respectively. Neglecting air resistance, after the projectile is fired it is acted upon by gravity only. Since the force of gravity acts vertically, the acceleration in the horizontal direction is 0 and the acceleration in the vertical direction is $g = -32$ (the acceleration due to gravity acting downward). Hence,

$$a_x = D_t v_x = 0 \quad \text{and} \quad a_y = D_t v_y = -32.$$

Therefore,

$$v_x = D_t^{-1}(0) = c_1 \quad \text{and} \quad v_y = D_t^{-1}(-32) = -32t + c_2.$$

Since the initial values of v_x and v_y are 693 and 400, respectively,

$$v_x = 693 \quad \text{and} \quad v_y = -32t + 400.$$

If X and Y represent the horizontal and vertical distances, respectively, which the projectile moves,

$$D_t X = v_x \quad \text{and} \quad D_t Y = v_y.$$

Therefore,

$$X = D_t^{-1}(693) = 693t + c_3,$$

$$Y = D_t^{-1}(-32t + 400) = -16t^2 + 400t + c_4.$$

If, as indicated in Fig. 7, the origin is placed at the starting point of the projectile's motion, $c_3 = c_4 = 0$, and

$$X = 693t, \quad Y = -16t^2 + 400t.$$

From analytical geometry, it is clear that these equations are parametric equations of a parabola. If we wish to do so we can obtain the Cartesian equation of the path, or *trajectory*, of the projectile by eliminating the parameter t . The result is

$$Y = -\frac{16}{(693)^2} X^2 + \frac{400}{693} X.$$

EXAMPLE 2. If a ball is thrown horizontally from a roof 100 ft high with the initial velocity 60 ft per sec, find (a) the formula for the horizontal distance moved in t sec; (b) the formula for the height after t sec; (c) the Cartesian equation of the path; and (d) when, where, and with what velocity the ball strikes the ground.

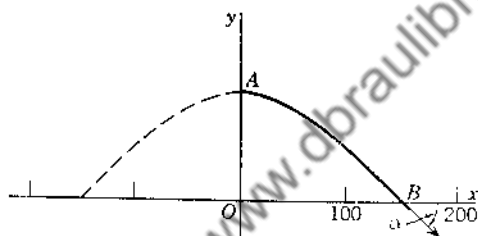


FIG. 8

SOLUTION. Since there is no horizontal acceleration, the horizontal velocity must be constantly 60 ft per sec. That is, $v_x = 60$ and $X = 60t + C_1$. If we are measuring all distances from a point on the ground directly under the point from which the ball is thrown, $X = 0$ when $t = 0$, and therefore $C_1 = 0$. Hence, $X = 60t$, which is the answer to (a).

The acceleration in the vertical direction is that due to gravity, which we shall assume to be -32 ft per sec per sec. Hence, $a_y = D_t v_y = -32$, and, consequently, $v_y = -32t + C_2$. Since the initial velocity in the vertical direction is 0, $C_2 = 0$, and $v_y = -32t$. Since $v_y = D_t Y$, we have $D_t Y = -32t$, whence $Y = -16t^2 + C_3$. Now, the ball is at height 100 when $t = 0$; hence, $C_3 = 100$ and $Y = -16t^2 + 100$. This is the answer to (b).

The two equations, $X = 60t$ and $Y = -16t^2 + 100$, are parametric equations of the path of the ball. To find the Cartesian equation, we eliminate t between these two equations and obtain $Y = -\frac{X^2}{225} + 100$.

This is the equation of the parabola shown in Fig. 8. Only the solid part of the curve represents the path of the ball, which we imagine thrown to the right from A and reaching the ground at B .

The ball strikes the ground when t has the positive value that makes $Y = 0$ in the equation $Y = -16t^2 + 100$. To find this value we solve the equation $-16t^2 + 100 = 0$, obtaining $t = 10/4$, or $2\frac{1}{2}$. The ball strikes the ground at B , to the right of A , the distance OB being the horizontal distance it moves in $2\frac{1}{2}$ sec. Putting $t = 2\frac{1}{2}$ in the equation $X = 60t$, we find the distance to be 150. The ball therefore strikes the ground 150 ft away from O , in $2\frac{1}{2}$ sec.

When the ball strikes the ground, the horizontal component of velocity is still 60 ft per sec. The vertical component is found by putting $2\frac{1}{2}$ in the equation $v_y = -32t$; hence at this time $v_y = -32 \cdot (2\frac{1}{2}) = -80$. The magnitude of the vector v is $|v| = \sqrt{(60)^2 + (-80)^2} = 100$. The direction of v is such that $\tan \alpha = 80/60 = 1.3333$, and $\alpha = 53^\circ$, approximately.

We cannot find the length of the path at this time, since application of formula 7.2 leads to

$$s = D^{-1}\sqrt{60^2 + 32^2t^2} = D^{-1}(4\sqrt{225 + 64t^2}),$$

and, as yet, we have no rule for evaluating this inverse derivative.

EXAMPLE 3. A particle moves according to the equations, $X = t^3/3$, $Y = t^2/2$, where t is the time in seconds and X and Y are measured in feet. (a) Find the expressions for v_x , v_y , a_x , and a_y . (b) Find the vectors v and a at the point $(\frac{1}{3}, \frac{1}{2})$. (c) Draw the path. (d) Find the length of the path from the start, at $t = 0$, until the particle reaches $(72, 18)$, at $t = 6$.

SOLUTION. (a) Since $v_x = D_t X$, $v_y = D_t Y$, $a_x = D_t v_x$, $a_y = D_t v_y$, we have

$$v_x = t^2, \quad v_y = t, \quad a_x = 2t, \quad a_y = 1. \quad (8.1)$$

(b) The point $(\frac{1}{3}, \frac{1}{2})$ is reached when $t = 1$. At this time, by substitution of $t = 1$ in Eqs. 8.1, we have $v_x = 1$, $v_y = 1$, $a_x = 2$, $a_y = 1$. Hence

$$|v| = \sqrt{1+1} = \sqrt{2},$$

and v is directed so that $\cos \alpha = 1/\sqrt{2}$, $\sin \alpha = 1/\sqrt{2}$; i.e., $\alpha = 45^\circ$. At the same time

$$|a| = \sqrt{4+1} = \sqrt{5},$$

and a is directed so that $\cos \beta = 2/\sqrt{5}$, $\sin \beta = 1/\sqrt{5}$. (c) The path can

be drawn directly from its parametric equations by means of the following table.

t	0	1	2	3	4	5	6
X	0	$1/3$	$8/3$	9	$64/3$	$125/3$	72
Y	0	$1/2$	2	$9/2$	8	$25/2$	18

A small part of the curve is shown in Fig. 9A, with the velocity vector at $(\frac{1}{3}, \frac{1}{2})$, and the same part is shown in Fig. 9B, with the acceleration vector

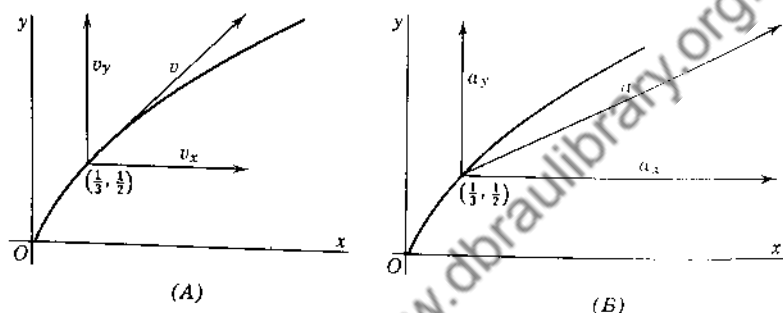


FIG. 9

at that point. (d) In order to find the length of the path, we use Eq. 7.2 with $\phi(t) = \sqrt{v_x^2 + v_y^2} = \sqrt{t^4 + t^2} = \sqrt{t^2 + 1} \cdot t$. That is,

$$s = D^{-1}(\sqrt{t^2 + 1} \cdot t).$$

By means of the general power formula I-1a, we obtain

$$s = \frac{1}{3}(t^2 + 1)^{\frac{3}{2}} + c. \quad (8.2)$$

Since the length is to be measured from the point (0, 0) that corresponds to $t = 0$, we have $s = 0$ when $t = 0$, i.e.,

$$0 = \frac{1}{3}(0 + 1)^{\frac{3}{2}} + c = \frac{1}{3} + c.$$

Hence $c = -\frac{1}{3}$. Substituting this value in (8.2), we obtain a formula for s measured from (0, 0) to a point corresponding to any value t .

$$s = \frac{1}{3}(t^2 + 1)^{\frac{3}{2}} - \frac{1}{3}. \quad (8.3)$$

Since the point (72, 18) is reached when $t = 6$, the required length is found by putting $t = 6$ in (8.3), and is $s = \frac{1}{3}(36 + 1)^{\frac{3}{2}} - \frac{1}{3} = 74.5$, approximately.

EXERCISES

1. A ball is thrown with initial velocity 50 ft per sec, in the direction $\alpha = \tan^{-1}(3/4)$. Assuming the origin to be at the point where the ball is thrown, find (a) parametric equations for its velocity; (b) parametric equations of its path; (c) the Cartesian equation of its path; (d) when, where, and with what velocity it strikes the ground; (e) the speed when the ball is at its highest point, i.e., $v_y = 0$.

2. The acceleration of a particle is $a_x = 4$, $a_y = 0$, and it starts from the origin with $|v| = 2$ and is directed vertically upward. Find (a) the acceleration vector, the velocity vector, and the position, when $t = 1$; (b) parametric equations of the path; and (c) draw the path.

3. A point moves away from the origin in the first quadrant along the parabola $y^2 = 2x$ in such a way that the X -coordinate increases at the rate of 2 units per sec. How fast does it pass through the point (2, 2)? [Hint: Differentiate $y^2 = 2x$ with respect to t . Thus, $2yD_t y = 2D_t x$, or $2yv_y = 2v_x$. Now, v_x is given equal to 2 and it is possible to find v_y at the point (2, 2). Hence, find $|v|$.]

4. A particle moves according to the equations $X = t^3$, $Y = t^2$. (a) Find expressions for v_x , v_y , a_x , and a_y . (b) Find the direction and magnitude of v when $t = 2$; (c) of a when $t = 2$. (d) Find the length of the path from (0, 0) to (8, 4).

5. A particle moves according to the equations $X = t^2 - 6t$, $Y = 2t$. (a) Find expressions for v_x , v_y , a_x , and a_y . (b) Find the position of the particle and the direction and magnitude of its velocity when $t = 7$; (c) of its acceleration when $t = 7$. (d) Draw the path.

6. A projectile is fired with an initial velocity of 1000 ft per sec in a direction inclined 30° to the horizontal. (a) Find parametric equations of the path, neglecting air resistance. (b) Find the horizontal range when fired at this inclination. (c) Find the greatest height that the projectile reaches.

7. A bomb is dropped from an airplane that is flying horizontally at the speed of 200 mph. (a) Find parametric equations of the path, neglecting air resistance. (b) Express the Cartesian equation of the path. (c) If the bomb is dropped from a height of 10,000 ft, find the time required for it to reach the surface of the earth, and the point of striking.

8. A body starts from rest at the origin, and in t seconds its acceleration parallel to the X -axis is 12, while the acceleration parallel to the Y -axis is $8t - 10$. Find parametric equations of the path. Find the minimum point of the path. [Hint: Find t for $v_y = 0$.]

9. Finding Curves from the Slope Formula. The inverse of the problem of finding the formula for the slope of the tangent to a given curve is that of finding the equation of the curve when given the formula for the slope. Let us consider some simple examples.

EXAMPLE 1. Find a curve having the slope always equal to $\frac{1}{2}$.

SOLUTION. From analytical geometry we know that the answer is a straight line which rises $\frac{1}{2}$ unit for each unit of horizontal advance. Moreover, any line rising in this way is a solution of the problem. The equation of the family of lines satisfying this condition is

$$y = \frac{1}{2}x + c, \quad (9.1)$$

where c can be any real number. Thus, there are infinitely many solutions.

Let us add a requirement by asking for a line whose slope is always equal to $\frac{1}{2}$ and which passes through $(0, 3)$. Since the equation has the form $y = \frac{1}{2}x + c$, this equation must be satisfied by $(0, 3)$. Hence,

$$3 = \frac{1}{2} \cdot 0 + c, \quad \text{and} \quad c = 3.$$

Consequently, the equation of the particular line called for is

$$y = \frac{1}{2} \cdot x + 3.$$

This problem can also be solved by the inverse derivative. We write

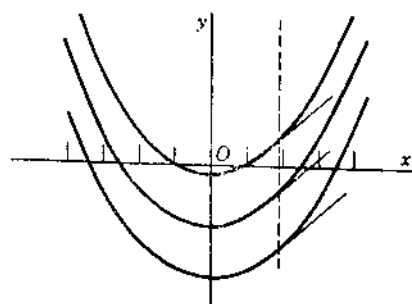


FIG. 10

SOLUTION. We write

$$Dy = \frac{1}{2}x,$$

whence

$$y = D^{-1}(\frac{1}{2}x),$$

or

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

As in Example 1, there are infinitely many curves which satisfy the condition of the problem, corresponding to the infinitely many values which c can assume. Several of these are shown in Fig. 10. If, in addition, we require that the curve pass through $(0, -3)$, we must have

$$-3 = \frac{1}{4}0^2 + c, \quad \text{or} \quad c = -3.$$

The equation is then

$$y = \frac{1}{4}x^2 - 3.$$

It is obvious geometrically from this example that the process of inverting differentiation ought to lead to infinitely many results, for, if one of the parabolas $y = \frac{1}{4}x^2 + c$ has its slope always equal to $\frac{1}{2}x$, so does any other in the family. The whole system may be thought of as generated by moving any one of them parallel to the Y -axis. At points of intersection with any line parallel to the Y -axis, all the slopes are equal.

$$\begin{aligned} Dy &= \frac{1}{2}x, \\ \text{whence} \quad y &= D^{-1}(\frac{1}{2}x), \\ \text{or} \quad y &= \frac{1}{4}x^2 + c. \end{aligned}$$

This equation agrees with (9.1), and the final solution can now be found as above.

EXAMPLE 2. Find a curve having its slope always equal to half the abscissa, and passing through $(0, -3)$.

For example, wherever one of these curves meets the line $x = 2$, the slope of the tangent is 1.

EXAMPLE 3. Find the equation of the family of curves having the slope of the tangent always equal to $3x^2 - 2x + 1$. Find the particular curve of the family that passes through (1, 3).

SOLUTION. We write

$$Dy = 3x^2 - 2x + 1$$

and

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

whence

$$y = x^3 - x^2 + x + c.$$

This is the required equation of the family. To find the particular curve that passes through (1, 3), we suppose that the equation is satisfied by $x = 1$, $y = 3$. Thus

$$3 = 1^3 - 1^2 + 1 + c, \quad \text{and} \quad c = 2.$$

Hence,

$$y = x^3 - x^2 + x + 2$$

is the equation of the particular curve required. The curves of the family corresponding to $c = -4, 0, 2$, and 5 are shown in Fig. 11.

In a similar way, we can find a curve for which the slope is given by any formula of the form $Dy = f(x)$, provided that we can carry out the inverse differentiation $D^{-1}f(x)$. After performing this operation, a value of the arbitrary constant can be found so that the curve will pass through any specified point, with some exceptions (see Exercises 11 and 12 below).

FIG. 11

EXERCISES

Find the equations of the systems of curves for which the slopes are given by the formulas in Exercises 1-10. With each, find the equation of the curve of the system that passes through the specified point, and draw the curve.

1. $Dy = \frac{1}{x}$; (2, 1).

2. $Dy = 2x$; (-2, 1).

3. $Dy = 2x - 1$; (1, 3).

4. $Dy = x^2$; (-3, -4).

5. $Dy = \sqrt{x}$; (1, 5).

6. $Dy = 2/x^2$; (-2, 1).

7. $Dy = 1/\sqrt{x}$; (4, 5).

8. $Dy = 1 + 1/\sqrt{x}$; (9, 16).

9. $Dy = (x - 1)^2$; (3, 6).

10. $Dy = (x - 1)(x + 1)$; (2, 2).

11. Show that no curve of the system in Exercise 6 passes through (0, 0); through (0, b) for any value of b.

12. Is there a curve of the system in Exercise 8 passing through (0, 0)? What can be said about the slope at that point?

10. Differential Equations. In our earliest work in mathematics we solved problems by arithmetic, working only with known numbers. Increased power was gained, when we studied algebra, by the use of letters to stand for unknown numbers. We set up equations containing these unknowns as well as known numbers and, by solving the equations, determined the unknowns.

We no longer limit our search to the values of unknown numbers. We have come to regard the knowledge of functional relations as a primary goal and the computation of particular values of the letters, or variables, that are involved as of less importance, often merely incidental. Previous to our work in calculus, our methods for finding functional relations to satisfy the requirements of problems were restricted to the processes of algebra and those introduced in the study of trigonometry and logarithms. After we learned to differentiate, functions were found by this process. A powerful new method of discovering functional relations has now been opened to us by the process discussed in this chapter. From the knowledge of the rate of change of a function we can find the function itself. Thus, we set up equations involving the derivative of a function, or the second derivative, or some higher derivative, and reason back to the function itself.

This enlargement of our methods is in some degree analogous to that brought about by the use of unknowns in algebra. We reasoned with the unknowns to obtain equations whose solutions gave us the values of the unknowns. We now set up equations involving unknown functions and reason from these equations in order to find the functions. Equations of this kind are called *functional equations* (see Chapter 1, Art. 4, Exercises 11 and 12).

The equations set up in this chapter involved the derivatives of the functions to be found. Such equations are called *differential equations*. Thus, in Art. 9, Example 2, we started with the knowledge that $Dy = \frac{1}{2}x$. We considered y to be a function of x . But at the start we did not know what that function was. We merely knew that its derivative was $\frac{1}{2}x$. $Dy = \frac{1}{2}x$ is an example of a differential equation. When we performed inverse differentiation to obtain $y = \frac{1}{4}x^2 + c$, we solved the differential equation. A differential equation and its solutions are defined as follows.

DEFINITION. A differential equation is an equation containing a derivative of an unknown function. A solution of a differential equation is a function that satisfies it when inserted in place of the unknown function.

To solve a differential equation completely we would have to find all its solutions. By virtue of the theorem in Art. 3 we know that $y = \frac{1}{4}x^2 + c$ gives all the solutions of the differential equation $Dy = \frac{1}{2}x$.

From a general expression for the result we are often interested in selecting a particular solution that satisfies some special condition. Thus, we found $y = \frac{1}{4}x^2 - 3$ to be the solution representing the special curve that we sought.

The power of the methods arising from the study of differential equations can be sensed only in a very minor degree at present, because of the limitations to our knowledge of differentiation and inverse differentiation. We have thus far learned how to apply these operations to rather simple algebraic functions only. We must extend them to more complicated algebraic functions, and to trigonometric, logarithmic, and exponential functions. We are further limited in that we have considered only differential equations of the form $D_x y = f(x)$. Although it may seem like a trivial extension to set the derivative equal to a function of x and y to get the form $D_x y = f(x, y)$, this is a very far-reaching extension indeed, and one that can be dealt with only in very simple examples in this book. Some examples that we shall study in later chapters are $D_x y = y$, $D_x^2 y = k^2 y$, and $D_x^2 y = -k^2 y$. Each of the last two examples is a differential equation of the second order, since it contains the second derivative of the unknown function. A differential equation as simple in appearance as $D_x y = x^2 + y^2$ offers difficulties that we shall avoid.

The exercises in Art. 3 and Art. 4 are really exercises in solving differential equations. The applications of inverse differentiation to problems in motion and to curves are applications of differential equations. Many more differential equations will be encountered as the uses of calculus operations are extended to other functions and other applications.

The Area Problem

1. Restatement of the Problem. In Chapter 1 it was pointed out that two geometric problems could serve as prototypes of the fundamental problems of calculus. One of them was the problem of finding the slope of the tangent to a curve. This was also termed the rate problem. We solved this problem for algebraic functions, and examined several of its applications and extensions. We now return to the second problem, the area problem. If we can solve this problem we shall find that a host of further applications can be dealt with successfully. Let us restate the problem.

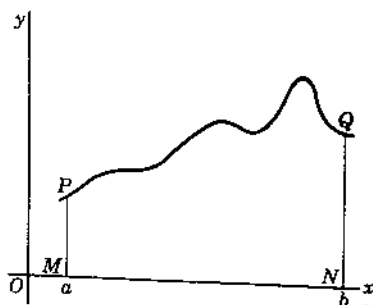


FIG. 1

We assume that a function $y = f(x)$ has for its graph, in the range $a \leq x \leq b$, a continuous curve, which we shall for the present suppose is entirely above the x -axis. Our problem is to find the area bounded by the curve, the x -axis, and the two ordinates $x = a$ and $x = b$ (Fig. 1).

It was pointed out in Chapter 1 that the elementary definition of area is meaningless for figures having curved boundaries, and, hence, cannot be applied in the present instance, at least not directly. But even though the area $MNQP$ in Fig. 1 is really not defined as yet, we are intuitively satisfied that this figure has an area. Furthermore, although we have no way of measuring this area exactly, we can make an estimate of its value. A method of making such an estimate was described in Chapter 1, and we saw there how we could go on to make better and better approximations to the area as conceived intuitively.

This was also the way we attacked the tangent problem, and we found that we could build up a sequence of ever-improving approximations. This sequence had a limit that we accepted as the solution of the problem and, at the same time, the definition of the slope of the tangent. We shall proceed in the same way with the area problem. First we shall see how to calculate approximations to some areas of this kind.

2. The Area under the Graph of $y = x$. The area of the trapezoid $MNQP$ in Fig. 2, bounded by the line $y = x$, the x -axis, and the ordinates $x = 5$ and $x = 10$, is easily found by elementary geometry to be $37\frac{1}{2}$. It will nevertheless be instructive to estimate the area by a method that can be applied to areas less easy to calculate. As a first approximation, which we call A_1 , we take the area of the rectangle $MNRP$. This is, of course, very inaccurate, but it will serve as a starting point in our sequence. We have

$$A_1 = \text{area } MNRP = 5 \cdot 5 = 25.$$

As a second approximation, we might take the sum of the five rectangles each with base = 1 and having the altitudes 5, 6, 7, 8, and 9, respectively. Thus

$$A_2 = 1 \cdot 5 + 1 \cdot 6 + 1 \cdot 7 + 1 \cdot 8 + 1 \cdot 9 = 35.$$

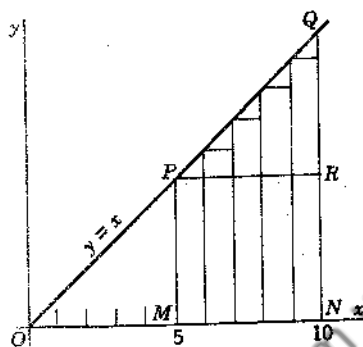


FIG. 2

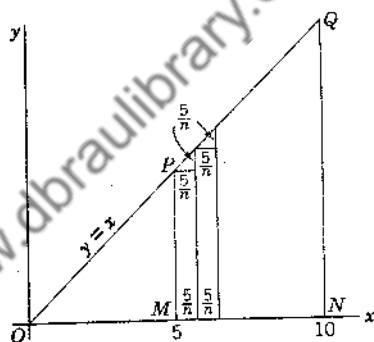


FIG. 3

We could form a third approximation by taking the sum of twenty-five rectangles, each with base = $\frac{1}{5}$ and altitudes 5, $5\frac{1}{5}$, $5\frac{2}{5}$, etc., the last one having the altitude $9\frac{4}{5}$. The student can show that

$$A_3 = (\frac{1}{5})(5 + 5\frac{1}{5} + 5\frac{2}{5} + \cdots + 9\frac{4}{5}) = 37.$$

Now, let us imagine the area approximated by n rectangles of equal width = $5/n$. The heights may be 5 , $5 + 5/n$, $5 + 2(5/n)$, $\cdots$, $5 + (n-1)(5/n)$, respectively, so that each rectangle just reaches up to the line $y = x$. We add $5/n$ to the height as we pass from one rectangle to the next because the height of the line $y = x$ increases at the same rate that x increases (Fig. 3). The value of A_n is therefore

$$A_n = 5(5/n) + (5 + 5/n)(5/n) + (5 + 2 \cdot 5/n)(5/n) + \cdots + (5 + [n-1] \cdot 5/n)(5/n)$$

$$= (5/n)\{5 + (5 + 5/n) + (5 + 2 \cdot 5/n) + \cdots + (5 + [n-1] \cdot 5/n)\}$$

$$= (5/n)\{5n + 5/n + 2 \cdot 5/n + \cdots + [n-1] \cdot 5/n\}$$

$$= 25 + (25/n^2)(1 + 2 + \cdots + [n-1]).$$

The quantity in parenthesis is an arithmetic progression, having the sum $(n^2 - n)/2$. Hence

$$\begin{aligned} A_n &= 25 + (25/n^2) \cdot (n^2 - n)/2 = 25 + 25n^2/2n^2 - 25n/2n^2 \\ &= 37\frac{1}{2} - 25/2n. \end{aligned}$$

Now, it is clear that $\lim_{n \rightarrow \infty} A_n = 37\frac{1}{2}$, so that we are led to conclude that we can obtain the exact area by taking the limit of the sequence of approximating sums, A_n .

3. The Symbol Σ . The calculation in Art. 2 required the summation of the progression $1 + 2 + 3 + \cdots + (n - 1)$. Similarly, the application of this method to the approximation of an area under any graph will require the summation of a progression of some kind. In Art. 4 we shall discover a method by which we can sum a good many series, including a series of the squares of the first n whole numbers, a series of their cubes, and a series of their k powers for any positive integral k . With this method we shall be in a position to approximate the area under a curve belonging to any polynomial.

The number of terms in these series, which is often n (an indefinite number), makes it advantageous to use a compact symbol for their sum. Such a symbol is provided by Σ , the Greek capital *sigma*. This letter is written in front of a formula for the successive terms, with an indication of the starting and terminating terms of the series. Thus

$$\sum_{i=1}^{i=10} i = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10.$$

$$\sum_{i=1}^{i=5} 1/i = 1/1 + 1/2 + 1/3 + 1/4 + 1/5.$$

$$\sum_{i=0}^{i=5} 3 \cdot 2^i = 3 \cdot 2^0 + 3 \cdot 2^1 + 3 \cdot 2^2 + 3 \cdot 2^3 + 3 \cdot 2^4 + 3 \cdot 2^5.$$

The formula for successive terms contains a letter that is understood to assume all integral values, in order, from that indicated at the bottom of the Σ to that indicated at the top. The particular letter used for the "running" variable, like i in these examples, is immaterial, since it does not appear in the result. It is common practice to abbreviate the denotations above and below the Σ , writing $\sum_{i=0}^5$ instead of $\sum_{i=0}^{i=5}$, or even $\sum_0^5$, if the identity of the running variable is clear. Thus, we might write

$$\sum_1^4 1/i^2 = 1/1^2 + 1/2^2 + 1/3^2 + 1/4^2.$$

Other examples are

$$\sum_{i=1}^{i=n} (1 + ih)^2 = (1 + h)^2 + (1 + 2h)^2 + \cdots + (1 + nh)^2.$$

$$\sum_{j=1}^n j/x_j = 1/x_1 + 2/x_2 + 3/x_3 + \cdots + n/x_n.$$

Since the number of terms symbolized by $\sum_1^n$ is finite, the associative, commutative, and distributive postulates of algebra apply. For example, $\sum_{i=1}^n (a_i + b_i) = \sum_{i=1}^n a_i + \sum_{i=1}^n b_i$, and $\sum_{i=1}^n ka_i = k \sum_{i=1}^n a_i$.

EXERCISES

Write in full and calculate Exercises 1-5.

$$1. \sum_1^4 i^2. \quad 2. \sum_1^4 i^3. \quad 3. \sum_1^5 (1 + 2i). \quad 4. \sum_1^4 i \cdot 2^i. \quad 5. \sum_1^5 1/i.$$

Write in full Exercises 6-9.

$$6. \sum_1^4 \sqrt{i}. \quad 7. \sum_0^5 1/(1 + i^2). \quad 8. \sum_{i=0}^5 f(i). \quad 9. \sum_{i=1}^5 if(i).$$

Write four terms and the last term in Exercises 10-12.

$$10. \sum_1^n i^2. \quad 11. \sum_0^{n-1} i^2. \quad 12. \sum_{i=1}^8 (1 + ih)h_i.$$

$$13. \text{ Show that } \sum_{i=1}^n 1 = n.$$

4. A Fundamental Summation Formula. Many series can be summed with the help of the following theorem.

$$\sum_{i=1}^n [f(i) - f(i-1)] = f(n) - f(0). \quad (4.1)$$

This formula is easily verified by writing the summation in full. Thus,

$$\begin{aligned} \sum_{i=1}^n [f(i) - f(i-1)] &= f(1) - f(0) + f(2) - f(1) + f(3) - f(2) \\ &\quad + \cdots + f(n) - f(n-1) \\ &= -f(0) + f(1) - f(1) + f(2) - f(2) \\ &\quad + \cdots - f(n-1) + f(n). \end{aligned}$$

Except for the first and last, the terms cancel in pairs, giving (4.1) as a result.

EXAMPLE 1. Calculate $\sum_{i=1}^{25} [\sqrt{i} - \sqrt{i-1}]$.

SOLUTION. Applying (4.1) with $f(i) = \sqrt{i}$, we get

$$\sum_{i=1}^{25} [\sqrt{i} - \sqrt{i-1}] = \sqrt{25} - \sqrt{0} = 5.$$

EXAMPLE 2. Calculate $\sum_{i=1}^{10} \frac{1}{(i+1)i}$.

SOLUTION. The fraction $\frac{1}{(i+1)i}$ can be expressed in the form $\frac{1}{i} - \frac{1}{i+1}$, as can be verified by combining this difference into one fraction. Therefore

$$\sum_{i=1}^{10} \frac{1}{(i+1)i} = \sum_{i=1}^{10} \left[\frac{1}{i} - \frac{1}{i+1} \right] = \sum_{i=1}^{10} \left[\left(-\frac{1}{i+1} \right) - \left(-\frac{1}{i} \right) \right].$$

Hence, letting $f(i) = -\frac{1}{i+1}$, we have

$$f(i-1) = -\frac{1}{i}, \quad f(10) = -\frac{1}{11}, \quad \text{and} \quad f(0) = -1.$$

Therefore, by formula 4.1,

$$\sum_{i=1}^{10} \frac{1}{(i+1)i} = -\frac{1}{11} - (-1) = 1 - \frac{1}{11} = \frac{10}{11}.$$

EXERCISES

1. Show that $\sum_{i=1}^n 1/(i+1)i = 1 - 1/(n+1)$.

2. Calculate $\sum_{i=1}^{25} 1/(\sqrt{i} + \sqrt{i-1})$. [Hint: Rationalize the denominator.]

3. Calculate $\sum_{i=1}^n (2i-2) = \sum_{i=1}^n [i(i-1) - (i-1)(i-2)]$.

4. Show that $\sum_{i=1}^{10} (3i^2 - 3i) = \sum_{i=1}^{10} [i(i+1)(i-1) - i(i-1)(i-2)]$
 $= 990;$

and express the value of $\sum_{i=1}^n (i^2 - i)$.

5. Show that $\sum_{i=1}^{20} 1/i(i+1)(i+2) = -1/2 \cdot \sum_{i=1}^{20} [1/(i+1)(i+2) - 1/i(i+1)]$, and hence calculate it. Express the value of $\sum_{i=1}^n 1/i(i+1)(i+2)$.

5. **Formulas for $\sum_1^n i^2$ and $\sum_1^n i^3$.** By means of the theorem embodied in formula 4.1 we can deduce formulas for sums of powers of the first n integers. Thus, since

$$i^3 - (i-1)^3 = 3i^2 - 3i + 1,$$

we have

$$\sum_1^n (3i^2 - 3i + 1) = \sum_1^n [i^3 - (i-1)^3] = n^3, \quad (5.1)$$

by (4.1). Now,

$$\sum_1^n (3i^2 - 3i + 1) = 3 \sum_1^n i^2 - 3 \sum_1^n i + \sum_1^n 1. \quad (5.2)$$

The sum of the arithmetic progression, $\sum_1^n i$, is $n(n+1)/2$, and $\sum_1^n 1 = n$ (see Exercise 13, Art. 3). Hence, substituting in (5.2),

$$\sum_1^n (3i^2 - 3i + 1) = 3 \sum_1^n i^2 - 3n(n+1)/2 + n. \quad (5.3)$$

Therefore, from (5.1) and (5.3),

$$3 \sum_1^n i^2 = n^3 + 3n(n+1)/2 - n = n(2n+1)(n+1)/2,$$

and

$$\sum_1^n i^2 = n(2n+1)(n+1)/6. \quad (5.4)$$

Similarly, we write

$$i^4 - (i-1)^4 = 4i^3 - 6i^2 + 4i - 1.$$

Hence,

$$\sum_1^n (4i^3 - 6i^2 + 4i - 1) = \sum_1^n [i^4 - (i-1)^4] = n^4, \quad (5.5)$$

by (4.1). Since

$$\sum_1^n (4i^3 - 6i^2 + 4i - 1) = 4 \sum_1^n i^3 - 6 \sum_1^n i^2 + 4 \sum_1^n i - \sum_1^n 1, \quad (5.6)$$

and we have already found that

$$\sum_1^n i^2 = n(2n+1)(n+1)/6, \quad \sum_1^n i = n(n+1)/2, \quad \sum_1^n 1 = n,$$

we have, from (5.6),

$$\sum_1^n (4i^3 - 6i^2 + 4i - 1) = 4 \sum_1^n i^3 - n(2n+1)(n+1) + 2n(n+1) - n. \quad (5.7)$$

Substituting from (5.5) in (5.7) and transposing,

$$\begin{aligned} 4 \sum_1^n i^3 &= n^4 + n(2n+1)(n+1) - 2n(n+1) - n \\ &= n^2(n+1)^2. \end{aligned}$$

Therefore,

$$\sum_1^n i^3 = n^2(n+1)^2/4. \quad (5.8)$$

By comparing this result with the formula $\sum_1^n i = n(n+1)/2$ we see that

$$\sum_1^n i^3 = \left(\sum_1^n i\right)^2.$$

The process by which these formulas for $\sum_1^n i^2$ and $\sum_1^n i^3$ were derived can be extended to find $\sum_1^n i^k$ for any positive integral k (see Exercise 1).

From (5.4) and (5.8) we can easily obtain formulas for $\sum_1^n (i-1)^2$ and $\sum_1^n (i-1)^3$. Thus,

$$\begin{aligned} \sum_1^n (i-1)^2 &= 0^2 + 1^2 + 2^2 + \cdots + (n-1)^2 \\ &= 1^2 + 2^2 + \cdots + (n-1)^2 + n^2 - n^2 \\ &= \left(\sum_1^n i^2\right) - n^2. \end{aligned}$$

Hence, from (5.4),

$$\sum_1^n (i-1)^2 = n(2n+1)(n+1)/6 - n^2 = n(2n-1)(n-1)/6. \quad (5.9)$$

Similarly, we can show that

$$\sum_1^n (i-1)^3 = n^2(n-1)^2/4. \quad (5.10)$$

In the applications to area problems formulas 5.9 and 5.10 are as useful as formulas 5.4 and 5.8.

EXERCISES

1. Use the method of this article to obtain a formula for $\sum_1^n i^4$. [Hint: Start with $i^5 - (i-1)^5$.]
2. Calculate $\sum_1^{10} i^2$ by means of formula 5.4. Check by termwise addition.
3. Calculate $\sum_1^6 i^3$ by means of formula 5.8. Check by termwise addition.

4. Calculate $\sum_{i=1}^7 (i-1)^2$ by means of formula 5.10. Check by termwise addition.
5. Derive formula 5.10.
6. Find the value of $\sum_{i=1}^8 (1 + i + i^2 + i^3)$.
7. Find the value of $\sum_{i=1}^6 (i^2 - 3i^2 + 6i + 2)$.
8. Express the value of the sum $\sum_{i=1}^n (2i^2 + i^2 - 8i - 1)$.

6. The Area under the Curve $y = x^2$. The method of approximating the area in Art. 2 was so successful that we were able to obtain the exact area as the limit of a sequence of approximations. We now apply a similar method to approximate the area bounded by the parabola $y = x^2$, the x -axis, and the ordinate $x = 1$ (Fig. 4).

We imagine the interval from 0 to 1 on the x -axis to be subdivided into a number of equal parts, say 8, so that the length of each subinterval is $\frac{1}{8}$. Ordinates drawn at the points of subdivision divide the area to be evaluated into strips. We imagine the area of each strip to be approximated by that of the rectangle of width $\frac{1}{8}$ and height equal to the least ordinate in the strip.

The sum of the areas of approximating rectangles is an approximation to the desired area. Now, the area of the first rectangle is 0 since the least ordinate in the strip is 0; the area of the second rectangle is $\frac{1}{8} \cdot (\frac{1}{8})^2$, since its width is $\frac{1}{8}$ and its height is $y = x^2 = (\frac{1}{8})^2$; the area of the third rectangle is $\frac{1}{8} \cdot (\frac{2}{8})^2$, its height being $y = x^2 = (\frac{2}{8})^2$; etc. The approximating sum, which we denote by A_8 , since there are 8 subdivisions, is

$$A_8 = \frac{1}{8} \cdot 0 + \frac{1}{8} \cdot (\frac{1}{8})^2 + \frac{1}{8} \cdot (\frac{2}{8})^2 + \cdots + \frac{1}{8} \cdot (\frac{7}{8})^2,$$

or

$$\begin{aligned} A_8 &= \sum_{i=1}^8 \frac{1}{8} \cdot [(i-1)/8]^2 = (1/8^3) \cdot \sum_{i=1}^8 (i-1)^2 \\ &= (1/8^3) \cdot 8(2 \cdot 8 - 1)(8 - 1)/6, \text{ by (5.9)} \\ &= 35/128. \end{aligned}$$

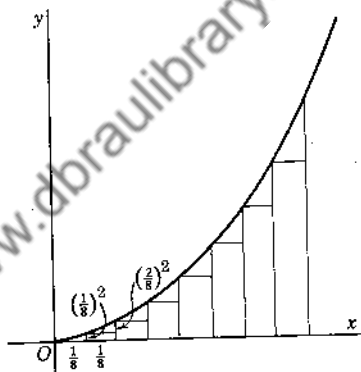


FIG. 4

Applying the same method to obtain an approximation by n rectangles of equal width, we obtain

$$\begin{aligned} A_n &= \sum_{i=1}^n \frac{1}{n} \cdot [(i-1)/n]^2 = (1/n^3) \cdot \sum_{i=1}^n (i-1)^2 \\ &= (1/n^3) \cdot n(2n-1)(n+1)/6, \text{ by (5.9)} \\ &= n(2n^2 - 3n + 1)/6n^3 = 1/3 - 1/2n + 1/6n^2. \end{aligned}$$

Since it is intuitively apparent that A_n approximates the area under the curve better and better as n increases, it is natural to define the area under the curve as $\lim_{n \rightarrow \infty} A_n$. By this definition, the area A is

$$A = \lim_{n \rightarrow \infty} A_n = \lim_{n \rightarrow \infty} (1/3 - 1/2n + 1/6n^2) = 1/3.$$

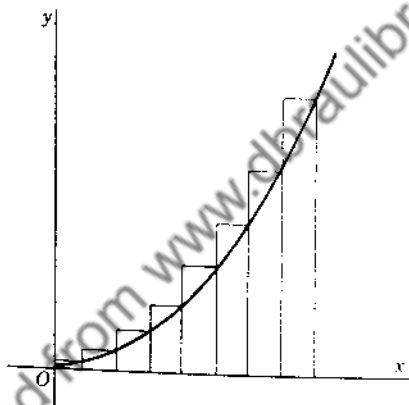


FIG. 5

EXERCISES

Apply the method of this article to find approximations A_n for each of the areas in Exercises 1-5. Find the area for each, assuming that the area $A = \lim_{n \rightarrow \infty} A_n$.

1. The area bounded by the curve $y = x^3$, the x -axis, and the ordinate $x = 1$. [Use formula 5.10.]
2. The area bounded by the parabola $y = 4x^2$, the x -axis, and the ordinate $x = 1$.
3. The area bounded by the parabola $y = x^2$, the x -axis, and the ordinate $x = 2$. [Hint: The width of each strip $\approx 2/n$.]
4. The area bounded by the curve $y = x^3/4$, the x -axis, and the ordinate $x = 1$.
5. The same area as in Exercise 4, with the ordinate $x = 1$ replaced by the ordinate $x = 2$.
6. Form an approximating sum for the area bounded by the curve $y = x^2$, the x -axis, and the ordinate $x = 1$ by means of rectangles drawn as in Fig. 5. Show that the limit of A_n formed this way is the same as that obtained above.

7. Definition of Area under a Curve. Returning to the problem stated in Art. 1, let us try to formulate a general definition of area under a curve (Fig. 6). We shall form a sequence of approximating sums, as in Art. 6. For the first approximation, we might take the area of the single rectangle

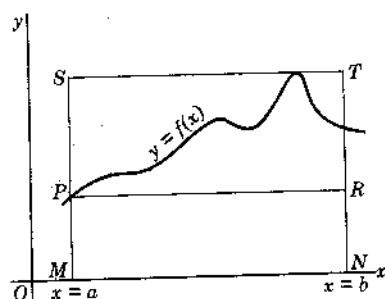


FIG. 6

$MNRP$, having the width $b-a$ and height equal to the least ordinate of the curve in the interval ab . We might equally well start with the area of the rectangle $MNTS$, having the width $b-a$ and height equal to the greatest ordinate in the interval. Neither of these approximations is

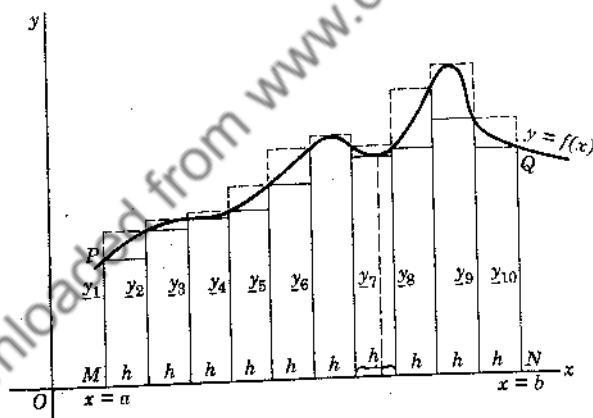


FIG. 7

good, but at least we feel sure that the area we seek lies between them. Thus, if we denote the area of $MNRP$ by S_1 , and that of $MNTS$ by $\bar{S}_1$, we state that $S_1 < A < \bar{S}_1$, where A represents the area sought.

For a second approximating sum we might divide the interval ab into 10 parts, each equal to $(b-a)/10$, which we denote by h , as in Fig. 7; at the points of subdivision draw ordinates forming 10 strips, each

bounded above by a segment of the curve; approximate the area of each strip just as we approximated the whole area in the previous step; and add the results. Let us denote the least ordinates in the various strips by $\underline{y}_1, \underline{y}_2, \dots, \underline{y}_{10}$. The sum $\underline{S}_2$ of the approximating rectangles that lie wholly below the curve is

$$\underline{S}_2 = \sum_{i=1}^{10} \underline{y}_i h = \left(\sum_{i=1}^{10} \underline{y}_i \right) h.$$

Forming, also, the sum $\overline{S}_2$ of rectangles having the same widths as before but with heights $\overline{y}_1, \overline{y}_2, \dots, \overline{y}_{10}$, each the greatest ordinate in its subdivision, we have

$$\overline{S}_2 = \sum_{i=1}^{10} \overline{y}_i h = \left(\sum_{i=1}^{10} \overline{y}_i \right) h.$$

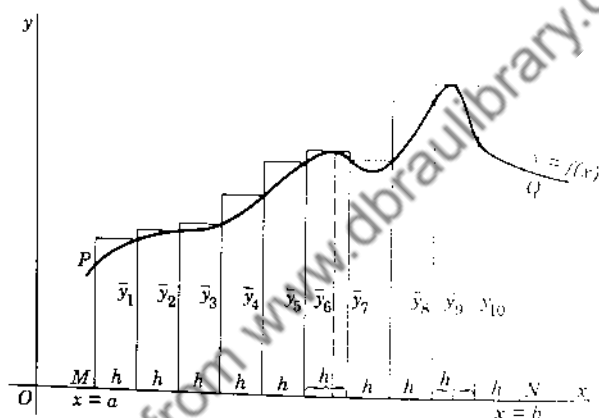


FIG. 8

$\overline{S}_2$ is the sum of the areas of the rectangles extending above the curve in Fig. 7. They are shown in Fig. 8 with their heights labeled.

As before, we are satisfied that $\underline{S}_2 < A < \overline{S}_2$. We observe also that $\underline{S}_1 < \underline{S}_2$ and $\overline{S}_1 > \overline{S}_2$. We have thus narrowed the range in which A can lie.

We might subdivide the interval ab still further and obtain $\underline{S}_3$ and $\overline{S}_3$ from rectangles corresponding to the new subdivisions, just as $\underline{S}_2$ and $\overline{S}_2$ were found. We could then write $\underline{S}_3 < A < \overline{S}_3$, and $\underline{S}_1 < \underline{S}_2 < \underline{S}_3$, $\overline{S}_1 > \overline{S}_2 > \overline{S}_3$.

In this manner we can form a sequence of lower approximating sums $\underline{S}_n$ and a sequence of upper approximating sums $\overline{S}_n$ with $\underline{S}_n < \overline{S}_n$. We would be convinced, also, that at each stage what we intuitively conceive the area to be would lie between $\underline{S}_n$ and $\overline{S}_n$.

If we could now show that the two sequences $\underline{S}_n$ and $\overline{S}_n$ both converge to the same limit, it would seem quite reasonable to accept this limit as the definition of the area A . This is, of course, what happened in Art. 6, although we did not exhibit the steps leading to the limit of the sequence of upper sums. This fact can be proved if $f(x)$ is any function that is continuous in the interval $a \leq x \leq b$. An understanding of the proof, however, requires some concepts and results that have not been presented. It will be found in more advanced textbooks.* The proof also shows the following facts, which are intuitively apparent.

1. The subdivisions of the interval MN do not have to be equal. In fact, any mode of subdivision can be used, provided only that the width of every subinterval tends to 0 as n tends to infinity. The same limit is reached irrespective of the mode of subdivision, as long as this condition is satisfied.

2. The same limit is obtained if each rectangle used in forming the sums S_n has any height lying between the least ordinate and the greatest ordinate in the interval. Of course, the sum S_n so formed would generally be different from $\underline{S}_n$ and $\overline{S}_n$, but the limit of the sequence of sums would be the same.

Assuming the truth of these statements without proof, for the present, we define the area bounded by the curve representing any positive continuous function, the x -axis, and two ordinates $x = a$ and $x = b$ with $a < b$, as follows.

1. Divide the interval ab into n subintervals in such a way that the length of each subinterval must tend to 0 as n tends to ∞ .

2. Form a rectangle in each subinterval with the length of the subinterval as width, and height equal to any ordinate of the curve in this subinterval.

3. Express the sum S_n of the area of the rectangles thus formed.†

4. Find the limit, $\lim_{n \rightarrow \infty} S_n$. The area is defined to be this result.

We have actually used this definition in Art. 6. The ideas of this article will be further illustrated in Art. 8.

8. The Area under the Curve $y = 1/x^2$. Let us try to find the area bounded by the curve $y = 1/x^2$, the x -axis, and the ordinates $x = 1$ and $x = 2$ (Fig. 9). The method we shall employ will closely parallel the process by which the definition of area was formulated in Art. 7. We

* See the statement of the theorem and footnote on p. 147.

† For simplicity of notation we here denote the sum of the areas of n rectangles by S_n ; in the illustration, S_2 was the sum of 10 rectangles. The number of rectangles involved in S_n is unimportant as long as it tends to ∞ with n .

shall form $\underline{S}_n$ and $\overline{S}_n$ and assume that $\lim_{n \rightarrow \infty} \underline{S}_n = \lim_{n \rightarrow \infty} \overline{S}_n$, in accordance with the theorem mentioned near the end of Art. 7. Inspection of these sums, however, will show how to form a third sum T_n , such that $\underline{S}_n < T_n < \overline{S}_n$, and it will be found that T_n has the constant value $\frac{1}{2}$. Hence, $\lim_{n \rightarrow \infty} \underline{S}_n = \lim_{n \rightarrow \infty} \overline{S}_n = \frac{1}{2}$.

First, we subdivide the interval from $x = 1$ to $x = 2$ into n equal parts, each having the length $h = 1/n$. The points of division are $1, 1+h, 1+2h, \dots, 1+(n-1)h, 1+nh = 2$. Let us form the sum of the

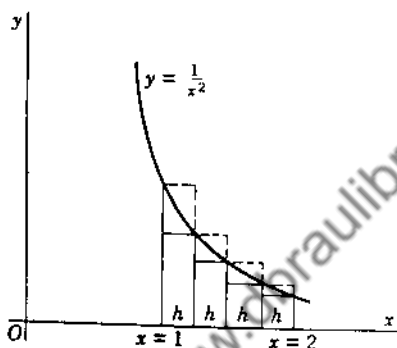


FIG. 9

areas of n rectangles corresponding to these subintervals, each having its height equal to the largest ordinate in its subinterval. In this example, the largest ordinate occurs at the beginning of each subinterval. Denoting the sum by $\overline{S}_n$, we have

$$\begin{aligned} \overline{S}_n = & h \cdot 1/1^2 + h \cdot 1/(1+h)^2 + h \cdot 1/(1+2h)^2 + \dots \\ & + h \cdot 1/(1+[n-1]h)^2. \end{aligned}$$

Since the least ordinate in each subinterval occurs at the right-hand end of the subinterval, we have for $\underline{S}_n$

$$\begin{aligned} \underline{S}_n = & h \cdot 1/(1+h)^2 + h \cdot 1/(1+2h)^2 + h \cdot 1/(1+3h)^2 + \dots \\ & + h \cdot 1/(1+nh)^2. \end{aligned}$$

We can now find a sum T_n that has its value between the values of $\underline{S}_n$ and $\overline{S}_n$ in the following way. Consider, for example, the second term in $\underline{S}_n$, which is $h/(1+2h)^2$, and the second term in $\overline{S}_n$, which is $h/(1+h)^2$. We have

$$h/(1+2h)^2 < h/(1+h)^2.$$

Now, if one of the factors $(1+2h)$ in the denominator of the first of these fractions is replaced by the smaller number $(1+h)$, the fraction

will be increased. But if only one of these factors is so replaced, the resulting fraction is still less than $h/(1+h)^2$. Hence

$$\frac{h}{(1+2h)^2} < \frac{h}{(1+2h)(1+h)} < \frac{h}{(1+h)^2}.$$

Let each of the terms of $\underline{S}_n$ be altered in this way, and denote the resulting sum by T_n .

$$T_n = \frac{h}{1 \cdot (1+h)} + \frac{h}{(1+h)(1+2h)} + \frac{h}{(1+2h)(1+3h)} + \cdots \\ + \frac{h}{(1+[n-1]h)(1+nh)}.$$

Then

$$\underline{S}_n < T_n < \overline{S}_n.$$

Now, each of the fractions in T_n can be broken into the difference of two fractions. Thus, for example,

$$\frac{h}{(1+h)(1+2h)} = \frac{1}{1+h} - \frac{1}{1+2h}.$$

Therefore

$$T_n = \left(\frac{1}{1} - \frac{1}{1+h} \right) + \left(\frac{1}{1+h} - \frac{1}{1+2h} \right) + \left(\frac{1}{1+2h} - \frac{1}{1+3h} \right) \\ + \cdots + \left(\frac{1}{1+[n-1]h} - \frac{1}{1+nh} \right).$$

Taking out the parentheses, all terms drop out except the first and last. The similarity of this step to the proof of formula 4.1 is apparent. For every value of n , therefore, this sum is equal to $1 - 1/(1+nh)$. Since $nh = 1$, we have $T_n = 1 - 1/(1+1) = \frac{1}{2}$. Now, since $\underline{S}_n$ and $\overline{S}_n$ converge to the same limit, and $\underline{S}_n < T_n < \overline{S}_n$, or $\underline{S}_n < \frac{1}{2} < \overline{S}_n$, this limit cannot be less than $\frac{1}{2}$ or more than $\frac{1}{2}$. It must therefore be exactly $\frac{1}{2}$. Hence, we conclude that the area in question is $\frac{1}{2}$ unit.

EXERCISES

1. Calculate the area bounded by the parabola $y = x^2$, the x -axis, and the ordinate $x = 1$, by using $\underline{S}_n$.
2. Find the area bounded by the curve $y = 1/x^2$, the x -axis, and the ordinates $x = 1$ and $x = 10$.
3. Find the area bounded by the curve $y = 1/x^2$, the x -axis, and the ordinates $x = a$ and $x = b$, if $0 < a < b$.

4. Find the area bounded by the curve $y = x + x^2$, the x -axis, and the ordinate $x = 1$. [Hint: Divide S_n into two parts S_n' and S_n'' , corresponding to the separate terms, x and x^2 .] Compare with the sum of the areas under $y = x$ and $y = x^2$.

5. Find the area bounded by the curve $y = x - x^2$, and the x -axis. [Hint: Note that the curve meets the x -axis at $x = 0$ and $x = 1$.] Show that this area is the same as that between the line $y = x$ and the parabola $y = x^2$.

6. Find the area bounded by the curve $y = x^2 - x^3$ and the x -axis. Show that this area is the same as that between the two curves $y = x^2$ and $y = x^3$.

7. Find the area bounded by the curve $y = x^2 + 1/x^2$, the x -axis, and the ordinate $x = 2$.

8. Find the area bounded by curve $y = x + 1/x^2$, the x -axis, and the ordinates $x = 1$ and $x = 2$.

9. The Definite Integral. Thus far in this chapter we have been pre-occupied with the geometric aspect of the area problem. Let us now

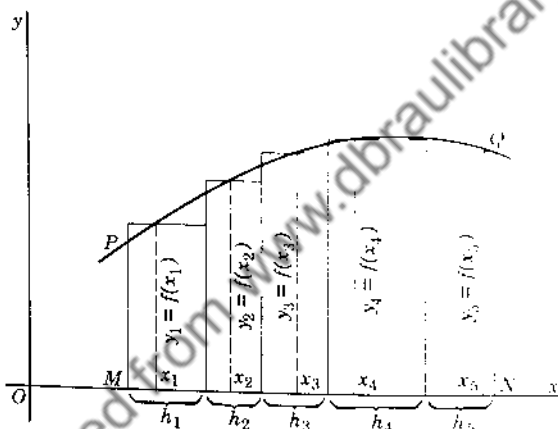


FIG. 10

examine more closely the numerical steps in the solution of the area problem.

To begin with, the upper boundary of any figure whose area has been under discussion is a curve that is the graphical representation of a numerical function, $y = f(x)$, over the interval $a \leq x \leq b$. The first step toward finding the numerical limit that we accept as the area is to subdivide the interval into n smaller intervals, n being any positive integer. The lengths of the subintervals may be represented by $h_1, h_2, \dots, h_n$, not necessarily all equal. They were taken equal in the examples, but this was a convenience rather than a necessity. The next step, looked on geometrically, consisted in building a rectangle over each subinterval with this subinterval as base and having its height equal to some ordinate of the curve over the subinterval. Numerically, this

means that in each subinterval we choose a value of x , say x_i , and find the corresponding value of $f(x)$, viz., $f(x_i)$. The geometrical step of expressing the area of each rectangle and adding the areas together is equivalent to forming the numerical products $f(x_i)h_i$ and adding them together.

Looked upon geometrically, what we now have is an approximation to the area under the curve, consisting of a sum of the areas of rectangles, each approximating the area of a vertical strip based on a subinterval of the x -axis. For example, the area $MNQP$ in Fig. 10 is approximated by the sum of five rectangles. Denoting the height of the i th rectangle by y_i , this sum is

$$S_5 = y_1h_1 + y_2h_2 + y_3h_3 + y_4h_4 + y_5h_5.$$

Since $y_1 = f(x_1)$, where x_1 is some value of x in the first subinterval, $y_2 = f(x_2)$, $y_3 = f(x_3)$, etc., we can write

$$S_5 = f(x_1)h_1 + f(x_2)h_2 + f(x_3)h_3 + f(x_4)h_4 + f(x_5)h_5.$$

Numerically, what we have is simply the sum of five products, each formed by multiplying the length of a subinterval by the functional value corresponding to some value of x in the subinterval.

In order to obtain an acceptable value of the area, we conceived the above process to be repeated over and over, with an ever increasing number of subintervals, thus forming an infinite sequence of approximating sums S_n . We required that this sequence be formed in such a way that as n tends to infinity the lengths of all subintervals tend to 0. The area was then defined to be

$$\text{Area} = \lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} (y_1h_1 + y_2h_2 + \cdots + y_nh_n).$$

This is, of course, the same thing as writing

$$\text{Area} = \lim_{n \rightarrow \infty} [f(x_1)h_1 + f(x_2)h_2 + \cdots + f(x_n)h_n].$$

Considered from the numerical point of view only, this limit is simply the limit of a sequence of sums of products formed from the function in the particular way described above. That is, it is

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i)h_i.$$

A further change in the representation of the sum is usually made. The letters $h_1, h_2, \cdots, h_n$, denote lengths of subintervals of the x -axis. Numerically, they are differences of values of x . It is usual to denote

such differences by $\Delta_1 x = h_1$, $\Delta_2 x = h_2$, etc.* (Δ is the Greek capital delta.) With this change,

$$f(x_1)h_1 + f(x_2)h_2 + \cdots + f(x_n)h_n = f(x_1)\Delta_1 x + f(x_2)\Delta_2 x + \cdots + f(x_n)\Delta_n x.$$

That is,

$$\sum_{i=1}^{i=n} f(x_i)h_i = \sum_{i=1}^{i=n} f(x_i)\Delta_i x.$$

The calculation of the limit, $\lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i)\Delta_i x$, is an application of the limit process to a sum involving the function $f(x)$ over the interval $a \leq x \leq b$. Just as in the operation of differentiation we signify the passing from the finite sum $\sum_{i=1}^n f(x_i)\Delta_i x$ to the limit of this sum as the number of terms tends to infinity, by a change in the symbol. Thus

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i)\Delta_i x \text{ is denoted by } \int_a^b f(x) dx.$$

In this symbol, n and i do not appear, since we no longer have the sum of n terms of the form $f(x_i)\Delta_i x$, but the *limit* of such a sum: the Greek letters Σ and Δ are replaced by the elongated S , $\int$, and the Roman d , respectively. The upper and lower limits of the interval of the variable x are attached to the symbol $\int$, making it $\int_a^b$. It is important to realize that the quantity expressed is no longer a sum in the familiar sense. It is called an *integral*. More precisely, $\int_a^b f(x) dx$ is called "the definite integral of $f(x)$ with respect to x from the limit a to the limit b ." The process of finding the integral is called *integration*. The numbers a and b are called the *limits of integration*. The term "limit" as applied here has the everyday meaning of "limit" as a "boundary," rather than that of a limit of a function or of a sequence. The function $f(x)$ to which the integration is applied is called the *integrand*.

We can now say that the area bounded by the curve $y = f(x)$, the x -axis, and the ordinates $x = a$ and $x = b$, is given by the integral

$$\int_a^b f(x) dx.$$

$$\text{Area} = \int_a^b f(x) dx.$$

* This is also often denoted by $\Delta x_1, \Delta x_2, \cdots, \Delta x_n, \cdots$. The notation $\Delta_i x$ is chosen in order to emphasize the enumeration of subintervals, e.g., the i th subinterval of values of x .

It is not possible to prove this statement. It follows from two definitions: the area is defined to be a certain limit, and the integral is defined to be the same limit.

We have not proved that $\sum_{i=1}^n f(x_i) \Delta_i x$ has a limit. This would, of course, be precisely the same as proving the existence of $\lim_{n \rightarrow \infty} S_n$ in Art. 7.

We state the theorem on the existence of this limit, for reference, without proof.*

Theorem on the existence of the integral of a continuous function. Let $f(x)$ be continuous in the interval $a \leq x \leq b$. Subdivide this interval into n subintervals having the lengths $\Delta_1 x, \Delta_2 x, \dots, \Delta_n x$, respectively, in such a way that the length of each subinterval must tend to 0 as n tends to ∞ . In each subinterval choose a value of x , say x_i in the i th subinterval, and form $f(x_1), f(x_2), \dots, f(x_n)$ and the sum $\sum_{i=1}^n f(x_i) \Delta_i x$. Then $\lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i) \Delta_i x$ exists, and is the same for all methods of subdivision that meet the above condition and for all choices of $x_1, x_2, \dots, x_n$, in the subintervals.

In applications we are led to set up sums like $\sum_{i=1}^n f(x_i) \cdot g(x'_i) \Delta_i x$, where x_i and x'_i are not necessarily the same but both lie in the interval $\Delta_i x$. The proof of the above theorem shows that this sum also tends to $\int_a^b f(x) \cdot g(x) dx$; that is, the limit is the same as though x_i and x'_i were the same. A similar statement holds for sums formed in other ways, e.g.,

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \sqrt{[f(x_i)]^2 + [g(x'_i)]^2} \Delta_i x = \int_a^b \sqrt{[f(x)]^2 + [g(x)]^2} dx.$$

We shall make use of these facts in Chapters 14, 15, and 16.

In finding the areas under the straight line, the parabola, and the curve $y = 1/x^2$, we really evaluated the integrals

$$\int_5^{10} x dx = 37\frac{1}{2}, \quad \int_0^1 x^2 dx = \frac{1}{3}, \quad \int_1^2 1/x^2 dx = \frac{1}{2}.$$

The student also evaluated several integrals in the exercises in Art. 6 and Art. 8.

EXAMPLE. Evaluate $\int_0^5 (x^2 - x) dx$ by direct application of the definition of the integral.

* For proof, see Courant, Vol. I, p. 131.

SOLUTION. We have $\int_0^5 (x^2 - x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n (x_i^2 - x_i) \Delta_i x$. Let the n subdivisions of the interval from 0 to 5 be equal so that $\Delta_i x = 5/n$, and let x_i be the greatest x in the i th subinterval. That is, $x_1 = 5/n$, $x_2 = 2 \cdot 5/n$, $x_3 = 3 \cdot 5/n$, $\dots$, $x_i = i \cdot 5/n$, $\dots$. Then

$$\begin{aligned} \sum_{i=1}^n (x_i^2 - x_i) \Delta_i x &= \sum_{i=1}^n [i \cdot 5/n]^2 - (i \cdot 5/n) \cdot 5/n \\ &= (5/n) \left[\sum_{i=1}^n (5/n)^2 i^2 - \sum_{i=1}^n (5/n) i \right] \\ &= (5/n)^3 \sum_{i=1}^n i^2 - (5/n)^2 \sum_{i=1}^n i \\ &= (125/n^3) \cdot [n(2n+1)(n+1)/6] \\ &\quad - (25/n^2) \cdot [n(n+1)/2] \\ &= 125/3 + 125/2n + 125/6n^2 - 25/2 - 25/2n. \end{aligned}$$

Hence

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n (x_i^2 - x_i) \Delta_i x = 125/3 - 25/2 = 175/6.$$

EXERCISES

Apply the definition of the integral to find the values of the integrals in Exercises 1-16.

- $\int_0^{10} 4 dx.$
- $\int_1^2 x dx.$
- $\int_1^2 (x+1) dx.$
- $\int_0^2 x^2 dx.$
- $\int_0^2 (x^2 + x) dx.$
- $\int_{-1}^0 x^2 dx.$
- $\int_{-3}^1 (x+3) dx.$
- $\int_{-1}^1 x^2/4 dx.$
- $\int_{-2}^2 x dx.$
- $\int_{-2}^1 x^3 dx.$
- $\int_0^2 x^3 dx.$
- $\int_{-2}^2 x^3 dx.$
- $\int_1^4 12/x^2 dx.$
- $\int_0^2 (x^4/4) dx.$
- $\int_0^4 (x^2 - 4x + 4) dx.$
- $\int_0^6 (x^3 - 4x) dx.$
- Show that $\int_a^b k dx = k(b-a).$
- Show that $\int_a^b kx dx = k(b^2 - a^2)/2.$
- Show that $\int_a^b kx^2 dx = k(b^3 - a^3)/3.$
- Show that $\int_a^b kx^3 dx = k(b^4 - a^4)/4.$

10. Theorems on Integrals. Even though we were led to the definition of a definite integral by consideration of the area under a curve, this definition is really independent of any geometric interpretation. It is a general numerical concept, one interpretation of which is the area under a curve. The restriction that $f(x)$ be positive throughout the interval $a \leq x \leq b$ was necessary only to avoid confusion concerning the notion of area. The requirement, $a < b$, was stipulated for the same reason.

The sum $\sum_{i=1}^n f(x_i) \Delta_i x$ can be formed equally well if $f(x)$ is positive, negative, or changes sign in the interval. It can be formed equally well whether $a < b$ or $a > b$. If $a > b$, the subdivision of the interval ab starts at the right-hand end of the interval and proceeds to the left, each $\Delta_i x$ being negative. Passing to the limit to obtain $\int_a^b f(x) dx$ from $\sum_{i=1}^n f(x_i) \Delta_i x$ is not affected by the removal of these restrictions. We shall therefore consider the definite integral to be defined for any function $f(x)$ that is continuous in the interval from a to b inclusive, where a and b are any two real numbers, with $a < b$, $a = b$, or $a > b$. If $a = b$, obviously $\int_a^b f(x) dx = 0$, since every subinterval is 0 and every sum $\sum_{i=1}^n f(x_i) \Delta_i x = 0$.

Some basic theorems concerning definite integrals are given below.

THEOREM 1. If $f(x)$ is the negative of $g(x)$ throughout the interval from a to b , i.e., $f(x) = -g(x)$, then $\int_a^b f(x) dx = -\int_a^b g(x) dx$.

This can be seen as follows. In forming the sums that have the limits $\int_a^b f(x) dx$ and $\int_a^b g(x) dx$, respectively, we might use $\sum_{i=1}^n f(x_i) \Delta_i x$ and $\sum_{i=1}^n g(x_i) \Delta_i x$, in which the same values are used for x_i and $\Delta_i x$ in the second sum as in the first. Since $f(x_i) = -g(x_i)$, it follows that

$$\sum_{i=1}^n f(x_i) \Delta_i x = \sum_{i=1}^n [-g(x_i)] \Delta_i x = -\sum_{i=1}^n g(x_i) \Delta_i x.$$

Taking the limits of these sums, we have

$$\int_a^b f(x) dx = -\int_a^b g(x) dx.$$

This theorem can also be expressed

$$\int_a^b f(x) dx = -\int_a^b [-f(x)] dx.$$

THEOREM 2. $\int_a^b f(x) dx = -\int_b^a f(x) dx$; that is, exchanging the limits changes the sign of the integral.

To see this we observe that, in the formation of the sums giving rise to $\int_b^a f(x) dx$, we can use the same terms that are used in forming the sums leading to $\int_a^b f(x) dx$, with $\Delta_i x$ in the one sum the negatives of $\Delta_i x$ in the other. Let us set up the corresponding sums, with $n = 4$, to see that this is so. In order to distinguish between the two, we replace x by x'

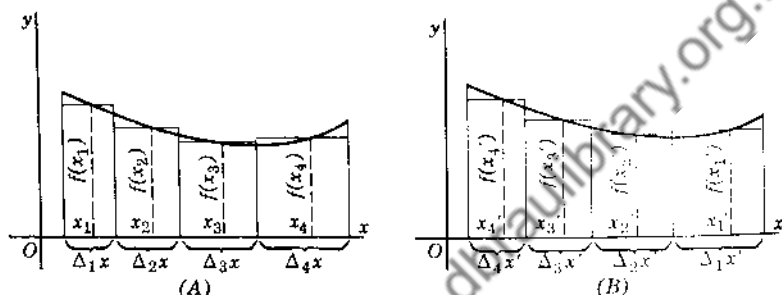


FIG. 11

in the second integral, writing it $\int_b^a f(x') dx'$, and writing the expression for the approximating sum: $\sum_{i=1}^n f(x'_i) \Delta_i x'$. We consider one such sum corresponding to the four strips shown in Fig. 11. The strips in (B) are identical with those in (A). Moreover, we can choose the values of x_i in the two sums to be in agreement in reverse order, as follows.

$$x'_1 = x_4, x'_2 = x_3, x'_3 = x_2, x'_4 = x_1.$$

If $a < b$, $\Delta_i x'$ is negative and $\Delta_i x$ is positive, so that $\Delta_1 x' = -\Delta_4 x$, $\Delta_2 x' = -\Delta_3 x$, $\Delta_3 x' = -\Delta_2 x$, $\Delta_4 x' = -\Delta_1 x$. It follows that

$$\sum_{i=1}^n f(x_i) \Delta_i x = -\sum_{i=1}^n f(x'_i) \Delta_i x'.$$

Since the limit of the sum on the left is $\int_a^b f(x) dx$, and that of the sum on the right is $\int_b^a f(x) dx$, we have

$$\int_a^b f(x) dx = -\int_b^a f(x) dx.$$

THEOREM 3. If $f(x)$ is continuous in the interval from a to c , where b lies between a and c , then

$$\int_a^b f(x) dx + \int_b^c f(x) dx = \int_a^c f(x) dx.$$

To see the truth of this statement, note that in the formation of each sum of the sequence of sums leading to $\int_a^c f(x) dx$ we can let one of the points of division be $x = b$. Then, for any n subintervals of the whole interval from a to c , let n_1 be the number of them that lie in the interval from a to b , and $n - n_1 = n_2$ the number of them that lie in the interval from b to c . We now form the sums $\sum_{i=1}^{n_1} f(x_i) \Delta_i x$, $\sum_{j=1}^{n_2} f(x_j) \Delta_j x$, and $\sum_{k=1}^n f(x_k) \Delta_k x$, corresponding respectively to the intervals from a to b , from b to c , and from a to c . Different running subscripts are used to help distinguish the intervals to which the sums apply. We now have

$$\sum_{i=1}^{n_1} f(x_i) \Delta_i x + \sum_{j=1}^{n_2} f(x_j) \Delta_j x = \sum_{k=1}^n f(x_k) \Delta_k x.$$

Consequently, on taking the limits as n_1 , n_2 , and n all tend to infinity so as to yield three integrals, we have the result stated in the theorem

THEOREM 4. $\int_a^b [f(x) + g(x)] dx = \int_a^b f(x) dx + \int_a^b g(x) dx$; that is, the integral of the sum of two functions is equal to the sum of their separate integrals over the same interval.

In order to show this, let us form $\sum_{i=1}^n f(x_i) \Delta_i x$ and $\sum_{i=1}^n g(x_i) \Delta_i x$ in such a way that the values of x_i in the two sums agree, as well as the values of $\Delta_i x$. We then have

$$\sum_{i=1}^n f(x_i) \Delta_i x + \sum_{i=1}^n g(x_i) \Delta_i x = \sum_{i=1}^n [f(x_i) + g(x_i)] \Delta_i x.$$

For example, if $n = 3$,

$$\begin{aligned} & f(x_1) \Delta_1 x + f(x_2) \Delta_2 x + f(x_3) \Delta_3 x + g(x_1) \Delta_1 x + g(x_2) \Delta_2 x + g(x_3) \Delta_3 x \\ &= [f(x_1) + g(x_1)] \Delta_1 x + [f(x_2) + g(x_2)] \Delta_2 x + [f(x_3) + g(x_3)] \Delta_3 x. \end{aligned}$$

Now, denoting $f(x) + g(x)$ by $\phi(x)$, this equation is

$$\sum_{i=1}^n f(x_i) \Delta_i x + \sum_{i=1}^n g(x_i) \Delta_i x = \sum_{i=1}^n \phi(x_i) \Delta_i x.$$

Taking the limits to obtain the integrals, we have

$$\int_a^b f(x) dx + \int_a^b g(x) dx = \int_a^b \phi(x) dx,$$

which is equivalent to the statement of the theorem.

THEOREM 5. $\int_a^b kf(x) dx = k \int_a^b f(x) dx.$

This follows from the equality

$$\sum_{i=1}^n kf(x_i)\Delta_i x = k \sum_{i=1}^n f(x_i)\Delta_i x,$$

whence

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n kf(x_i)\Delta_i x = k \left[\lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i)\Delta_i x \right].$$

EXAMPLE 1. Evaluate $\int_1^5 (-x^2) dx.$

SOLUTION. By Exercise 19, Art. 9, $\int_1^5 x^2 dx = (125 - 1)/3 = 124/3.$
Hence, by Theorem 1, $\int_1^5 (-x^2) dx = -124/3.$

EXAMPLE 2. Evaluate $\int_3^1 (x^2 + 2x) dx.$

SOLUTION. By Theorem 4,

$$\int_3^1 (x^2 + 2x) dx = \int_3^1 x^2 dx + \int_3^1 2x dx.$$

Now, by Theorem 2, $\int_3^1 x^2 dx = -\int_1^3 x^2 dx$, and $\int_3^1 2x dx = -\int_1^3 2x dx$,
and by Exercises 18 and 19, Art. 9,

$$\int_1^3 x^2 dx = (27 - 1)/3, \quad \int_1^3 2x dx = 2(9 - 1)/2.$$

Hence

$$\int_3^1 (x^2 + 2x) dx = -26/3 - 8 = -50/3.$$

EXAMPLE 3. Evaluate $\int_0^{10} f(x) dx$ if

$$f(x) = x, \text{ for } 0 \leq x \leq 5$$

$$= 14x - x^2 - 40, \text{ for } 5 < x \leq 10.$$

SOLUTION. Although this function is expressed by two different formulas in the two parts of the interval, it is continuous, since, at the point where the interval is divided,

$$f(x) = x = 5 \quad \text{and} \quad \lim_{x \rightarrow 5} (14x - x^2 - 40) = 5.$$

Therefore, the theorems on integrals apply. We make use of Theorem 3 and write

$$\int_0^{10} f(x) dx = \int_0^5 x dx + \int_5^{10} (14x - x^2 - 40) dx.$$

Now

$$\int_0^5 x dx = 25/2, \text{ by Exercise 18, Art. 10,}$$

$$\begin{aligned} \int_5^{10} (14x - x^2 - 40) dx &= \int_5^{10} 14x dx - \int_5^{10} x^2 dx - \int_5^{10} 40 dx, \text{ by Theorems} \\ &\quad \text{4 and 5} \\ &= 14(100 - 25)/2 - (1000 - 125)/3 - 40(10 - 5), \\ &\quad \text{by Exercises 17-19, Art. 9} \\ &= 33\frac{1}{3}. \end{aligned}$$

By addition,

$$\int_0^{10} f(x) dx = 25/2 + 33\frac{1}{3} = 12\frac{1}{2} + 33\frac{1}{3} = 45\frac{5}{6}.$$

EXERCISES

Evaluate the following integrals by use of the theorems of this article. The answers to exercises in Art. 9 can be employed wherever appropriate.

1. $\int_2^1 x dx.$
2. $\int_0^{-1} x^2 dx.$
3. $\int_1^{-3} (x + 3) dx.$
4. $\int_0^2 (x^3 + x^2 + 1) dx.$
5. $\int_0^2 (4x^3 - 2x^2 + 4x) dx.$
6. $\int_1^4 -(12/x^4) dx.$
7. $\int_1^4 (12/x^3 - 3x^2) dx.$
8. $\int_1^{-1} (x^4/4) dx.$
9. $\int_3^{15} \sqrt{2}(x + 2) dx.$
10. $\int_{-1}^{-3} (-x^2) dx.$
11. $\int_0^2 f(x) dx$, if $f(x) = x$, for $0 \leq x \leq 1$,
 $= 2x - x^2$, for $1 < x \leq 2$. Interpret the result as an area.

12. $\int_0^9 f(x) dx$, if $f(x) = 1$, for $0 \leq x \leq 1$,
 $= 1/x^2$, for $1 < x \leq 2$,
 $= 9/4 - x$, for $2 < x \leq 9/4$. Interpret the result as an area.

11. Need for a Better Method of Evaluating Integrals. Thus far we have been able to integrate only a small variety of functions. In Art. 6 and 8 we applied the definition of area to find the areas under some simple graphs. In Art. 9 the same process was used, this time under the name of integration, and without reference to the geometric problem of area. Some additional functions were dealt with in Art. 10, by using the theorems of that article, although these new functions were simple combinations of those that had been integrated earlier. It would be possible to find the integrals of a considerably greater variety of functions by direct application of the definition of the integral. This would be tedious, however, and far less productive than the program that will be followed in Chapter 8. There we shall find that an exceedingly important theorem, so important that it is called the fundamental theorem of calculus, will provide a method of integrating a great variety of functions that would be difficult or impossible to deal with at the present juncture.

Downloaded from www.digitallibrary.org

Evaluation of Integrals

1. A Mean Value Theorem. As a first step toward a new method of evaluating integrals we give our attention to a theorem that may be recognized as a restatement of a fact that was grasped intuitively at the start of our consideration of the area problem. This is called the *mean value theorem*, or, more accurately, the first mean value theorem for integrals, since there are other mean value theorems.

This theorem states that the integral of the continuous function $f(x)$ between the limits a and b is equal to the product of some value of the function in that interval multiplied by the length of the interval $b - a$. That is,

$$\int_a^b f(x) dx = f(x')(b - a), \quad (1.1)$$

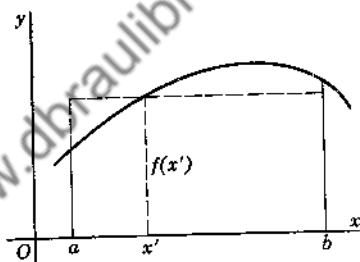


FIG. 1

where x' is some value of x in the interval $a \leq x \leq b$. Geometrically, this means that if $f(x)$ is positive throughout the interval, the area under the curve is equal to the area of a rectangle with base $= b - a$ and height equal to some value assumed by $f(x)$ in the interval (Fig. 1).

Although this theorem is intuitively obvious from geometrical considerations, as a theorem about the integral, which is a numerical quantity, it requires proof. First, we state that if M is the greatest value of $f(x)$ in the interval,

$$\int_a^b f(x) dx \leq \int_a^b M dx = M(b - a).$$

This is seen from the sums by which the integrals are defined. We compare $\sum_{i=1}^n f(x_i) \Delta_i x$ with $\sum_{i=1}^n M \Delta_i x$. Since $f(x_i) \leq M$ for all values of x_i , the first of these sums is certainly not greater than the second. The same must be true of the limits of the sums (see Exercise 7, Art. 6, Chapter 2).

In the same way, it follows that if m is the least value of $f(x)$ in the interval from a to b ,

$$m(b-a) \leq \int_a^b m \, dx \leq \int_a^b f(x) \, dx.$$

Hence,

$$m(b-a) \leq \int_a^b f(x) \, dx \leq M(b-a).$$

Since $\int_a^b f(x) \, dx$ is a definite number not less than $m(b-a)$ and not greater than another number $M(b-a)$, it must be equal to some number that can be written $\mu(b-a)$, where μ (Greek mu) is a number between m and M . Thus,

$$\int_a^b f(x) \, dx = \mu(b-a), \text{ where } m \leq \mu \leq M. \quad (1.2)$$

From Theorem 5, Art. 11, Chapter 2, we know that a function that is continuous throughout an interval assumes all values between its least value and its greatest value in that interval. Hence, we can set $\mu = f(x')$, where x' is some value of x in the interval (there may be more than one value of x in the interval for which $f(x) = \mu$). Replacing μ by $f(x')$ in (1.2),

$$\int_a^b f(x) \, dx = f(x')(b-a). \quad (1.1)$$

EXAMPLE. Find a value x' satisfying the mean value theorem if $a = 3$, $b = 6$, and $f(x) = x^2$.

SOLUTION. We find that $\int_3^6 x^2 \, dx = 63$. But, by the mean value theorem, $\int_3^6 x^2 \, dx = (x')^2(6-3)$, where x' is some value of x between 3 and 6. Thus,

$$3(x')^2 = 63, \quad \text{and} \quad x' = \sqrt{21}.$$

EXERCISES

Find a value of x' satisfying the mean value theorem in each of the following exercises.

1. $a = 0, b = 4, f(x) = x.$

2. $a = 3, b = 6, f(x) = x.$

3. $a = 0, b = 1, f(x) = x^2.$

4. $a = 1, b = 2, f(x) = \frac{1}{x^2}.$

5. $a = 2, b = 5, f(x) = x^2 - 2x + 3.$

2. A Function Defined by an Integral. Thus far we have considered integrals over definite intervals determined by fixed limits. In fact, the term "definite integral" means an integral over a definite interval. We

may extend the integral concept to indefinite intervals by supposing that one of the limits is variable. Thus, suppose the upper limit is not the fixed number b , but a variable u . Then it would be expected that the value of $\int_a^u f(x) dx$ would depend upon the value of u , and so be a function

of u : $\int_a^u f(x) dx = \phi(u)$. For example, replacing b by u in Exercises 17 and 18 of Art. 9, Chapter 7, we obtain $\int_a^u k dx = k(u - a)$ and $\int_a^u kx dx = \frac{k}{2}(u^2 - a^2)$.

Associating this idea with the area problem, we think of the area partially bounded by the graph of $y = f(x)$, the x -axis, and a fixed ordinate $x = a$ (Fig. 2). The figure is open on the right-hand side. If we bound it on this side by some ordinate $x = u_1$, we shall have a definite area. If we choose a different ordinate $x = u_2$, we shall have a different area. With all boundaries fixed except the ordinate $x = u$, the value of the area is dependent upon the location of this ordinate, i.e., upon u , and we can write $A = \phi(u)$.

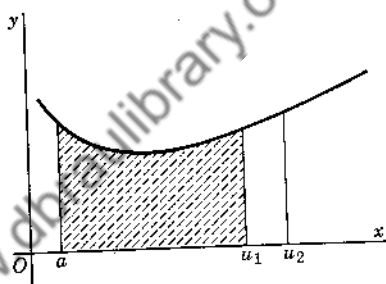


FIG. 2

The integral, $\int_a^u f(x) dx$, is called an *indefinite integral* of $f(x)$. By changing the lower limit we obtain other indefinite integrals of the same function. It should be observed, however, that any two of the indefinite integrals of the same function differ by a constant. Thus, if we replace a by a different lower limit c , we obtain the function $\int_c^u f(x) dx = \psi(u)$. Now, by Theorem 3, Art. 10, Chapter 7,

$$\int_a^u f(x) dx = \int_a^c f(x) dx + \int_c^u f(x) dx.$$

Since $\int_a^c f(x) dx$ is a definite number, say K , it follows that

$$\phi(u) = K + \psi(u).$$

As illustrated in Fig. 2, the values of u appear on the x -axis. It is a mistake, however, to identify u with x , since these variables play quite different roles. For each value of u there is an integration of $f(x)$; that is, we must think of the formation of a sequence of sums of the products $f(x_i)\Delta x$ and the determination of the limit of this sequence. Of course,

it is immaterial what letters are assigned to these variables; but to avoid confusion they should be different letters. Thus,

$$\int_a^u f(x) dx = \phi(u), \quad \int_a^t f(s) ds = \phi(t), \text{ etc.}$$

EXERCISES

Find the functions of u defined by the following integrals:

$$1. \int_1^u x dx.$$

$$2. \int_0^u (x - 3) dx.$$

$$3. \int_0^u t^2 dt.$$

$$4. \int_1^u \frac{1}{x^2} dx.$$

$$5. \int_0^u (3x^2 + x) dx.$$

$$6. \int_1^u (3x^2 + x) dx.$$

3. The Fundamental Theorem of Calculus. The key which we seek for the practical problem of calculating integrals is provided by the following theorem, which, on account of its great importance throughout calculus, we shall call the *fundamental theorem of calculus*.

FUNDAMENTAL THEOREM. The derived function of the function defined by an indefinite integral of a continuous function $f(x)$ has the same form as $f(x)$; that is, if

$$\phi(u) = \int_a^u f(x) dx,$$

$$\phi'(u) = f(u).$$

This fact has no doubt been evident in examples of indefinite integrals in Art. 2, as well as in the exercises of that article. Thus

$$\int_a^u kx dx = \frac{k}{2} (u^2 - a^2) = \phi(u),$$

and

$$\phi'(u) = ku.$$

Here $f(x) = kx$ and $f(u) = ku = \phi'(u)$.

In order to prove this theorem we differentiate $\phi(u)$ by direct application of the definition of a derivative. Thus, for any value u , we form

$$\phi(u + h) = \int_a^{u+h} f(x) dx.$$

Subtracting $\phi(u)$,

$$\begin{aligned} \phi(u + h) - \phi(u) &= \int_a^{u+h} f(x) dx - \int_a^u f(x) dx \\ &= \int_u^{u+h} f(x) dx \end{aligned}$$

by Theorem 3, Art. 10, Chapter 7.

Applying the mean value theorem,

$$\int_u^{u+h} f(x) dx = hf(x'),$$

where x' lies between u and $u + h$. Thus,

$$\phi(u + h) - \phi(u) = hf(x').$$

Dividing by h ,

$$\frac{\phi(u + h) - \phi(u)}{h} = f(x').$$

Taking the limit of this quotient as h tends to 0, in order to obtain the derivative,

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

Now, as $h \rightarrow 0$, $u + h \rightarrow u$, and, hence, x' tends to u . Therefore, since $f(x)$ is continuous at $x = u$,

$$\lim_{h \rightarrow 0} f(x') = f(u).$$

Hence,

$$\phi'(u) = f(u).$$

4. Application to the Evaluation of Integrals. The fundamental theorem tells us that an indefinite integral $\int_a^u f(x) dx$ is an inverse derivative $D^{-1}f(u)$, or primitive function of $f(u)$. It appeared in Chapter 6 that any two primitives of a function differ only by a constant. This will be proved in Chapter 12. Thus, if $F(u)$ is any single primitive function of $f(u)$,

$$\int_a^u f(x) dx = F(u) + c.$$

Suppose, now, that we have to evaluate the definite integral $\int_a^b f(x) dx$ for some particular function such as $f(x) = x^3$. We may write

$$\int_a^u x^3 dx = u^4/4 + c.$$

Now, $\int_a^b x^3 dx$ must be the value of this function when $u = b$. Thus,

$$\int_a^b x^3 dx = b^4/4 + c.$$

In order to evaluate c , we take $u = a$, obtaining

$$\int_a^a x^3 dx = a^4/4 + c.$$

Since

$$\int_a^a x^3 dx = 0,$$

we have

$$c = -a^4/4,$$

and

$$\int_a^b x^3 dx = b^4/4 - a^4/4.$$

In general, to evaluate $\int_a^b f(x) dx$, we first express

$$\int_a^u f(x) dx = F(u) + c,$$

where $F(u)$ is some primitive function of $f(u)$. We then replace u by b , obtaining

$$\int_a^b f(x) dx = F(b) + c.$$

In order to evaluate c , we replace u by a :

$$0 = \int_a^a f(x) dx = F(a) + c,$$

whence

$$c = -F(a).$$

Therefore

$$\int_a^b f(x) dx = F(b) - F(a). \quad (4.1)$$

Equation 4.1 may be considered a formula for evaluating definite integrals. It must be remembered, of course, that $F(u)$ is a primitive function of $f(u)$. In applying this formula we do not have actually to insert an unknown constant c and then evaluate it by setting $u = a$. Instead, we follow the rule:

RULE. Write a primitive function of $f(u)$ (or $f(x)$, since the name of the variable no longer matters), insert b in place of u in this function, then insert a in place of u , and subtract the results. The arbitrary constant can be omitted in the expression for the primitive function.

EXAMPLE 1. Evaluate $\int_a^b x^n dx$, if $n \neq -1$.

SOLUTION. $D^{-1}x^n = \frac{x^{n+1}}{n+1}$. Following the above rule,

$$\begin{aligned}\int_a^b x^n dx &= \frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1} \\ &= \frac{1}{n+1} (b^{n+1} - a^{n+1}).\end{aligned}$$

The form

$$\int_a^b f(x) dx = \left[D^{-1}f(x) \right]_a^b$$

is often convenient. The expression on the right-hand side stands for the inverse derivative of $f(x)$, first with b replacing x , and then with a replacing x , and the difference taken. Thus,

$$\int_a^b x^n dx = \left[\frac{x^{n+1}}{n+1} \right]_a^b = \frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}$$

EXAMPLE 2. Evaluate $\int_3^5 \sqrt{x^2 - 9} \cdot x dx$.

SOLUTION. By the method of Art. 4, Chapter 6,

$$D^{-1}(\sqrt{x^2 - 9} \cdot x) = \frac{1}{3}(x^2 - 9)^{\frac{3}{2}}.$$

Hence, by the rule stated above,

$$\begin{aligned}\int_3^5 \sqrt{x^2 - 9} \cdot x dx &= \left[\frac{1}{3}(x^2 - 9)^{\frac{3}{2}} \right]_3^5 = \frac{1}{3}(25 - 9)^{\frac{3}{2}} - \frac{1}{3}(9 - 9)^{\frac{3}{2}} \\ &= \frac{1}{3}(16)^{\frac{3}{2}} = \frac{64}{3}.\end{aligned}$$

EXAMPLE 3. Evaluate $\int_1^2 (3x^2 + 1/x^2) dx$.

SOLUTION. We have

$$D^{-1}(3x^2 + 1/x^2) = x^3 - 1/x.$$

Hence

$$\int_1^2 (3x^2 + 1/x^2) dx = \left[(x^3 - 1/x) \right]_1^2 = (8 - \frac{1}{2}) - (1 - 1) = 7\frac{1}{2}.$$

EXAMPLE 4. Find the area bounded by the curve $y = \sqrt[3]{x}$, the x -axis, and the ordinates $x = 1$ and $x = 8$.

SOLUTION. Expressing the required area by a definite integral,

$$A = \int_1^8 \sqrt[3]{x} dx.$$

Now, we have

$$D^{-1} \sqrt[3]{x} = \frac{3}{4} x^{\frac{4}{3}}.$$

Hence,

$$\begin{aligned} \int_1^8 \sqrt[3]{x} \, dx &= \frac{3}{4} (8)^{\frac{4}{3}} - \frac{3}{4} (1)^{\frac{4}{3}} \\ &= \frac{3}{4} \cdot 16 - \frac{3}{4} = 11\frac{1}{4}. \end{aligned}$$

The area is therefore $11\frac{1}{4}$ units.

EXERCISES

Evaluate the definite integrals in Exercises 1-9 by the method of this article.

1. $\int_0^{10} x \, dx.$

2. $\int_{-1}^1 3 \, dx.$

3. $\int_{-1}^1 (x^2 + 1) \, dx.$

4. $\int_1^{12} \frac{1}{x^2} \, dx.$

5. $\int_{\frac{1}{4}}^{\frac{1}{2}} \frac{1}{x^3} \, dx.$

6. $\int_{\frac{1}{4}}^{\frac{1}{2}} \left(2x^2 + \frac{1}{x^2} \right) \, dx.$

7. $\int_4^5 \sqrt{25 - x^2} \, dx.$

8. $\int_{-5}^5 \sqrt{25 - x^2} \, dx.$

9. $\int_0^2 \frac{x}{\sqrt{x^2 + 1}} \, dx.$

10. Interpret Exercise 2 as an area.

11. Interpret Exercise 4 as an area.

12. Interpret Exercise 5 as an area.

13. Calculate the area bounded by the curves $y = 4\sqrt{x}$ and $y = x^2/2$. [Hint: The points of intersection furnish the limits. Evaluate two integrals and subtract.]

14. Calculate the area bounded by the curves $y = x^2$ and $y = 8 - x^2$.

5. Primitive Functions and Integrals. At first acquaintance the tangent problem and the area problem exhibit no obvious relationship. The solutions of these problems seem to lead in quite different directions. The tangent problem is solved by the limit of a quotient. The values of the independent variable that are used in determining this limit are clustered about the point of tangency. The area problem is solved by the limit of a sum, and the values of the independent variable used in finding this limit are spread from beginning to end of the interval over which the area lies. It is, therefore, somewhat surprising to discover the close relationship between these problems.

It is fortunate that this relationship exists, for without it we would be much hampered in the solution of problems in finding areas and many other quantities that are most naturally approached as limits of sums leading to definite integrals. Because integrals are usually evaluated by means of inverse differentiation, it is customary to use the integral sign to symbolize the latter operation. Thus, $D^{-1}f(x)$ is usually written $\int f(x) \, dx$, with no limits attached to the integral sign. This expression is

often called the *indefinite integral* of $f(x)$. We shall reserve the term "indefinite integral," however, for expressions like $\int_a^u f(x) dx$. Although we shall often refer to $D^{-1}f(x) = \int f(x) dx$ as an integral, it must be remembered that it is an inverse derivative or primitive function of $f(x)$. With the integral notation the formulas and exercises of Chapter 6 would be written with D^{-1} replaced by $\int \cdots dx$. Thus, formula I-1 of Art. 3, Chapter 6, can be written

$$\int x^n dx = x^{n+1}/(n+1) + C, \quad \text{if } n \neq -1. \quad (I-1)$$

EXERCISES

Find the following primitive functions.

- $\int (5x^4 - 8x^3 + 6x^2 + 2x - 4) dx.$
- $\int (3x^2 - 2x + 4 - 4/x^2) dx.$
- $\int (1/\sqrt{x} + 5x^{\frac{3}{2}}) dx.$
- $\int (2x^2 - 3)^{\frac{1}{2}} x dx.$
- $\int (ax^3 - bx + c)(3ax^2 - b) dx.$
- $\int 2x\sqrt{x^2 + 1} dx.$
- $\int 2x/\sqrt{x^2 + 1} dx.$
- $\int (3x^3 + 8)^{\frac{1}{2}} x^2 dx.$

Some Applications of Integrals

1. Introduction. In this chapter we shall study some of the many applications of the definite integral. Here we shall limit our considerations to areas of plane figures, some special volumes, work done by a variable force, and the total force due to the pressure of a liquid on a

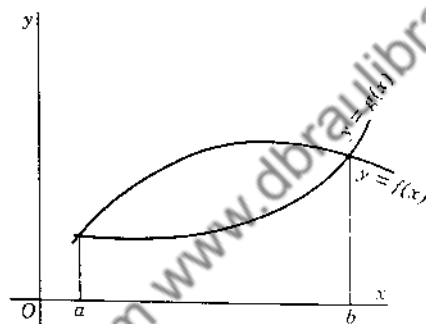


FIG. 1

vertical wall. Other applications will be taken up in later chapters. The present study of these applications is introductory and illustrative. The same applications will be treated more fully later on when we shall have more technique at our command.

2. Areas. It was the area problem which led us to the formulation of the definite integral. This problem therefore comes to our attention before other applications of the integral. We have already seen how the integral gives the area of a figure, one boundary of which is the graph of the positive function $y = f(x)$, the other boundaries being the x -axis and two ordinates $x = a$ and $x = b$. Here, if $a < b$,

$$\text{Area} = \int_a^b f(x) dx.$$

It remains for us to consider other plane figures with curved boundaries. In Exercises 13 and 14 of Art. 4, Chapter 8, one other example was

encountered, namely that in which the figure is completely bounded by the graphs of two positive functions, $y = f(x)$ and $y = g(x)$, Fig. 1. The obvious method of dealing with this example is first to solve for the intersections of $y = f(x)$ and $y = g(x)$ in order to find the limits, a and b , of the definite integrals to be applied, and then to evaluate the two definite integrals $\int_a^b f(x) dx$ and $\int_a^b g(x) dx$ and subtract. A slight variation of this method results from using Theorem 4, Art. 10, Chapter 7. Thus, since

$$\int_a^b f(x) dx - \int_a^b g(x) dx = \int_a^b [f(x) - g(x)] dx,$$

$$\text{Area} = \int_a^b [f(x) - g(x)] dx.$$

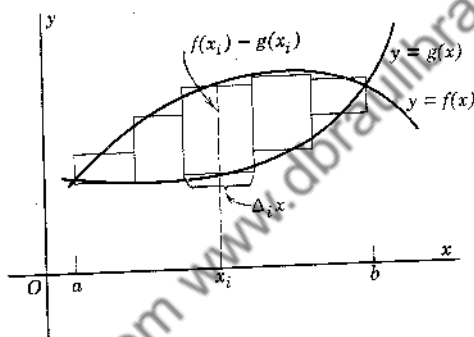


FIG. 2

The subtraction of $g(x)$ from $f(x)$ before the integration amounts to the same thing as setting up an integral as the limit of the sum

$$\sum_{i=1}^n [f(x_i) - g(x_i)] \Delta_i x.$$

Geometrically, this means that we obtain the area as the limit of a sequence of approximations, each approximating sum being formed by adding together the areas of rectangles drawn as indicated in Fig. 2. The approximating rectangle having the width $\Delta_i x$ has its height equal to the vertical distance between the two curves at some point x_i in the subinterval of width $\Delta_i x$.

When the boundary curve is the graph of a negative function $y = \phi(x)$, it is clear that the area has the numerical value of $\int_a^b \phi(x) dx$, Fig. 3.

But since this integral is negative in value, when $a < b$, and it is customary to think of an area as essentially positive, we write

$$\text{Area} = -\int_a^b \phi(x) dx, \quad \text{or} \quad \text{area} = \int_a^b [-\phi(x)] dx. \quad (1.1)$$

We may also write,

$$\text{Area} = \left| \int_a^b \phi(x) dx \right|. \quad (2.2)$$

Equation 2.2 is correct whether $\phi(x)$ is positive or negative. It is not correct if $\phi(x)$ changes sign in the interval, as in Fig. 4. In this case the

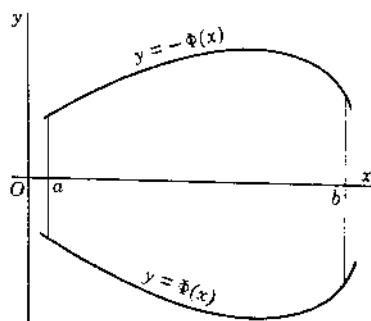


FIG. 3

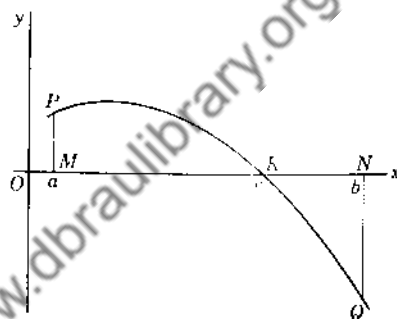


FIG. 4

number of units of area bounded by the curve $y = \phi(x)$, the x -axis, and the two ordinates $x = a$ and $x = b$, is equal to the area of MKP + the area of KQN . We have

$$\text{Area of } MKP = \int_a^c \phi(x) dx$$

$$\text{Area of } KQN = -\int_c^b \phi(x) dx$$

Thus, the total area

$$= \int_a^c \phi(x) dx - \int_c^b \phi(x) dx$$

or

$$= \left| \int_a^c \phi(x) dx \right| + \left| \int_c^b \phi(x) dx \right|.$$

No change in our formulas is required by areas which extend to the left-hand side of the y -axis. The above formulas were given, however, with the assumption that $a < b$. If $a > b$ there is, of course, a reversal of sign in accordance with Theorem 2, Art. 10, Chapter 7. One should

be particularly careful about signs when either or both of the numbers a and b are negative. Thus, in Fig. 5,

$$\text{Area } MNQP = \int_{-10}^{-5} f(x) dx,$$

$$\text{Area } HKSR = \int_{-1}^4 g(x) dx.$$

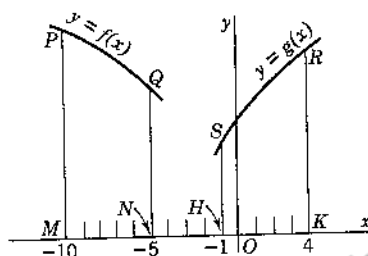


FIG. 5

EXAMPLE 1. Find the area bounded by the curve $y = x^2 - 6x - 3$, the x -axis, and the ordinates $x = 1$ and $x = 4$, Fig. 6.

SOLUTION. Area = $\left| \int_1^4 (x^2 - 6x - 3) dx \right| = |-33| = 33$.

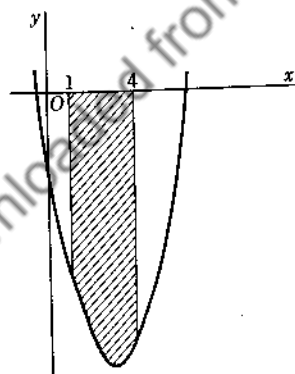


FIG. 6

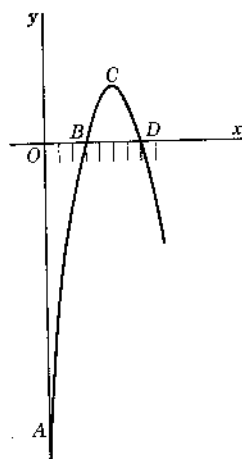


FIG. 7

EXAMPLE 2. Find the whole area bounded by the curve $y = 10x - x^2$ and the coordinate axes.

SOLUTION. From the curve (Fig. 7) we see that there are two parts of the area, OAB and BCD . The points B and D have ordinates equal to

0, and therefore their abscissas satisfy the equation $10x - x^2 - 21 = 0$. The two roots of this equation are $x = 3$ and $x = 7$. Hence, the abscissa of B is 3 and that of D is 7. The required area is now

$$\begin{aligned} \text{Area } OAB + \text{area } BCD &= \left| \int_0^3 (10x - x^2 - 21) dx \right| \\ &\quad + \left| \int_3^7 (10x - x^2 - 21) dx \right| \\ &= |-27| + |10\frac{8}{3}| = 37\frac{2}{3}. \end{aligned}$$

EXAMPLE 3. Calculate the area bounded by the curve $y = x^3 + 5x^2 + x + 5$, the x -axis, and the ordinate $x = 2$.

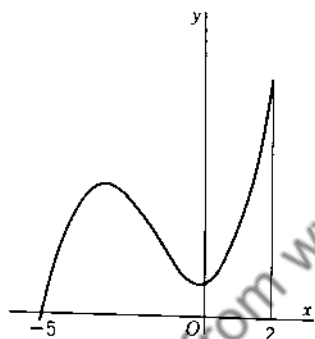


FIG. 8

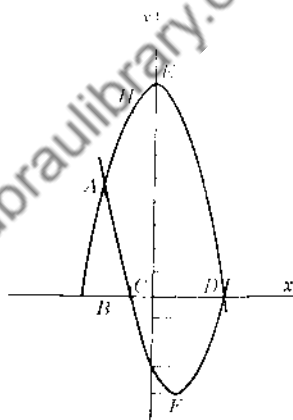


FIG. 9

SOLUTION. Factoring, $y = (x + 5)(x^2 + 1)$. Hence, the curve crosses the x -axis at $x = -5$, and at no other point since the factor $x^2 + 1$ is positive for all real values of x . The graph can be drawn by means of a short table of values (Fig. 8). Some assistance is given by locating the maximum and minimum points by the methods of Chapter 5. We find a maximum point at $x = -3.2$, $y = 20.3$, and a minimum point at $x = -0.1$, $y = 4.9$, all these coordinates being correct to one decimal place.

The required area is

$$\text{Area} = \int_{-5}^2 (x^3 + 5x^2 + x + 5) dx = 93\frac{11}{12}.$$

EXAMPLE 4. Calculate the area bounded by the parabolas $y = 9 - x^2$ and $y = x^2 - 2x - 3$, Fig. 9.

SOLUTION. Our first inclination is to find the areas of $ACDE$ and CFD and add the results. This is correct, of course. In order to find the area $ACDE$ we would first find area ACH by integrating $\int_{-2}^{-1} (9 - x^2) dx$ and subtracting $\int_{-2}^{-1} (x^2 - 2x - 3) dx$. We could then find the area $HCDE$ as the value of $\int_{-1}^3 (9 - x^2) dx$. The sum of the results for area ACH and area $HCDE$ is area $ACDE$. Now the area CFD is equal to $-\int_{-1}^3 (x^2 - 2x - 3) dx$. Adding the area of CFD to that of $ACDE$ we have

$$\begin{aligned} \text{Area of } ACFDE = & \int_{-2}^{-1} (9 - x^2) dx - \int_{-2}^{-1} (x^2 - 2x - 3) dx \\ & + \int_{-1}^3 (9 - x^2) dx - \int_{-1}^3 (x^2 - 2x - 3) dx. \end{aligned}$$

The sum of the first and third of these integrals is

$$\int_{-2}^{-1} (9 - x^2) dx + \int_{-1}^3 (9 - x^2) dx = \int_{-2}^3 (9 - x^2) dx,$$

by Theorem 3, Art. 10, Chapter 7. Similarly,

$$-\int_{-2}^{-1} (x^2 - 2x - 3) dx - \int_{-1}^3 (x^2 - 2x - 3) dx = -\int_{-2}^3 (x^2 - 2x - 3) dx.$$

Hence,

$$\begin{aligned} \text{Area } ACFDE &= \int_{-2}^3 (9 - x^2) dx - \int_{-2}^3 (x^2 - 2x - 3) dx \\ &= \int_{-2}^3 [(9 - x^2) - (x^2 - 2x - 3)] dx \\ &= \int_{-2}^3 (12 + 2x - 2x^2) dx \\ &= 41\frac{2}{3}. \end{aligned}$$

As illustrated by Fig. 2, we can think of this area as the limit of the sum of rectangles formed as indicated in Fig. 10. Here, the height of each rectangle is $f(x_i) - g(x_i)$, where $f(x) = 9 - x^2$ and $g(x) = x^2 - 2x - 3$. This is seen at once for the rectangles to the left of $x = -1$. For those to the right of $x = -1$, note that $g(x_i)$ is negative and, hence, $f(x_i) - g(x_i)$ is an arithmetic sum—the sum of the lengths of the ordinates $f(x_i)$ and $g(x_i)$.

EXAMPLE 5. Find the area bounded by the curve $y^2 = x^3$, the y -axis, and the line $y = 8$, Fig. 11.

SOLUTION. No new method is necessary to determine this area. Here we let y be the independent variable. We think of the interval of the y -axis from $y = 0$ to $y = 8$; subdivide it into n subintervals, form a rectangle in each subinterval approximating a horizontal strip of the area required, and take the limit of the sum of the areas of all these rectangles. This limit is the integral

$$\int_0^8 f(y) dy = \int_0^8 \sqrt[3]{y^2} dy = 19\frac{1}{3}.$$

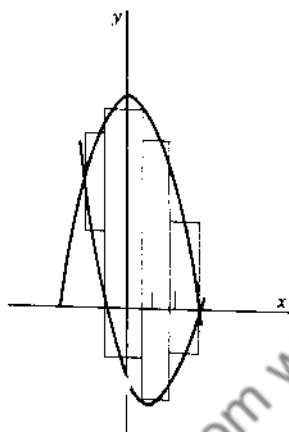


FIG. 10

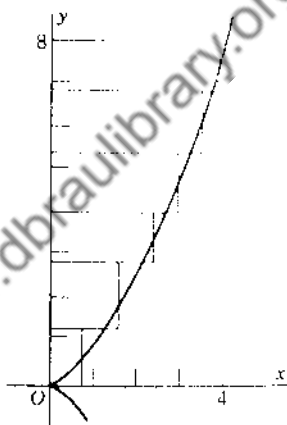


FIG. 11

EXERCISES

Find the areas bounded by the graphs of the following equations, the x -axis, and the ordinates specified.

- $y = (x - 2)^2 + 2$; $x = 0$, $x = 6$.
- $y = (x - 2)^2$; $x = 0$, $x = 6$.
- $y = (x - 2)^2 - 4$; $x = 0$, $x = 6$.
- $y = 1 - (x - 2)^2$; $x = 0$, $x = 6$.
- $y = x^3 - x$; $x = -1$, $x = 0$.
- $y = x^3 - x$; $x = -1$, $x = 1$.
- $y = x^3 - x$; $x = 0$, $x = 2$.
- $y = \sqrt{x}$; $x = 0$, $x = 4$.
- $y = -\sqrt{x}$; $x = 0$, $x = 4$.
- $y = x^{\frac{1}{2}}$; $x = -8$, $x = 8$.
- $y = 1/\sqrt{x}$; $x = \frac{1}{4}$, $x = 4$.
- $y = 1/x^{\frac{1}{2}}$; $x = -8$, $x = -1$.
- $y = x^{\frac{1}{2}}$; $x = -8$, $x = -1$.
- $y = 1/\sqrt{x} - 1/x^2$; $x = 1$, $x = 9$.
- Find the area bounded by the parabola $x^2 = 2y$ and the straight line $y = x + 4$.
- Find the area bounded by the parabola $y = 6x - x^2$ and the straight line $y = x - 6$.

18. Find the area between the curve $y = 1/x^2$ and the chord joining $(\frac{1}{2}, 4)$ and $(2, \frac{1}{4})$.

19. Find the area bounded by the curve $\sqrt{x} + \sqrt{y} = \sqrt{a}$ and the coordinate axes. (The curve is a part of a parabola tangent to both axes.)

20. Find the area bounded by the curve $x = \frac{1-y^2}{y^2}$ and the lines $x-1=0$, $y=1$, and $y=4$.

21. Find the area bounded by the curve $y = x^3 - 9x^2 + 24x - 7$ and the line $y=9$.

22. Find the area bounded by the curve $y^3 = x^2$ and the line $y=1$.

23. The rectangle $OACB$ in Fig. 12 is divided into two parts by the curve $y = x^{10}$. Find the ratio of the area of the part above the curve to that below the curve.

24. Solve Exercise 23, replacing the equation $y = x^{10}$ by $y = x^n$.

25. Find the area bounded by a parabola and its latus rectum in terms of the focal distance p .

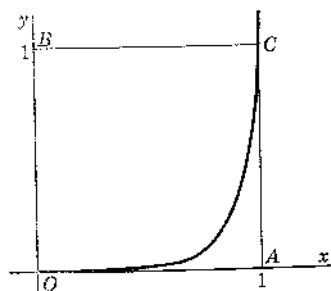


FIG. 12

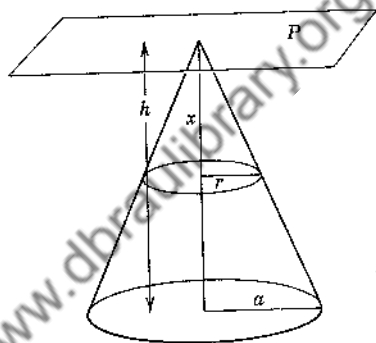


FIG. 13

3. Volumes of Some Special Solids. The volume of a solid bounded by planes is defined in solid geometry. For example, the volume of a rectangular parallelepiped is defined to be the number of unit cubes that can be fitted into it exactly. A unit cube is simply any cube we wish to choose as our standard. The number that represents the volume can be any positive real number, integral, fractional, or irrational. If any boundary is curved, of course this definition can no longer apply, since there is no way that cubes can be fitted exactly to such a solid. After our experience with areas it would be natural to try to define the volume of a solid figure having a curved bounding surface by means of an integral. We shall do this in the present article for some solids with special shapes.

We shall consider solids that meet the following condition: there is some plane such that every cross section parallel to this plane is a figure whose area can be expressed as a function of the distance of the cross section from the plane. For example, all circular cones satisfy this condition. Thus, in Fig. 13, every cross section of the cone parallel to

the base, and therefore parallel to the plane P through the vertex, which is itself parallel to the base, has the area $\pi(ax/h)^2$, where x is the distance of the cross section from plane P .

Another example is the figure pictured in Fig. 14. This figure might be approximated physically by placing two semicircular hoops as shown, meeting at right angles at P which is the midpoint of both semicircular arcs, and then winding a string tightly about the figure, starting at A and going around through B, C, D , and so on, making a surface one strand deep. Every cross section of the solid so bounded, parallel to the base $ABCD$, is a square, and its area can be expressed as a function of its distance

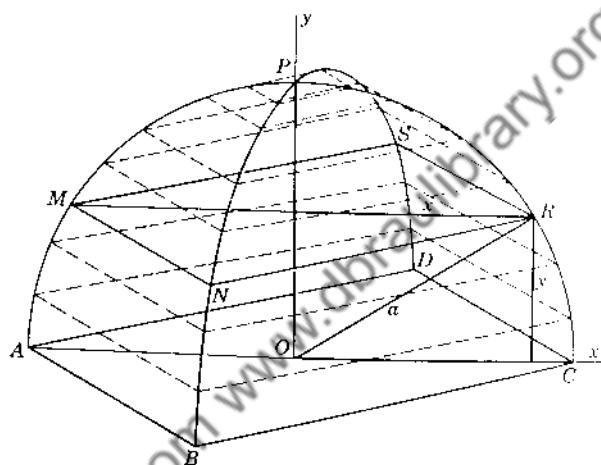


FIG. 14

from the base as follows. Suppose the semicircular hoops have the radius a . Then the equation of the circle of which APC is a part is $x^2 + y^2 = a^2$. Now, the area of a square is equal to one-half the square of its diagonal. Hence, the area of $MNRS = \frac{1}{2}(2x)^2 = 2x^2$. Since x is the abscissa of the point R on the arc APC , $x^2 = a^2 - y^2$. Therefore, the area of $MNRS$ is equal to $2(a^2 - y^2)$. Thus, the area of the cross section is a function of the distance y from the base plane.

An important class of figures satisfying the above condition are called *figures of revolution*. The right circular cone is one of these. A figure of revolution is conceived as being generated by revolving some plane figure about an axis. Thus, as the triangle OAB in Fig. 15 revolves about the line OA , it generates the cone. As the triangle sweeps around, every one of its points (except those on the axis OA) traces, or generates, a circle, and the totality of all the points of all circles so traced make up the cone. Similarly, Fig. 16 pictures the solid of revolution generated by revolving

about the x -axis the area OAP bounded by the parabola $y = \frac{1}{16}x^2$, the x -axis, and the line $x = 4$.

We shall take up the calculation of volumes of solids of revolution in Art. 4. Volumes of other solids meeting the condition laid down above,

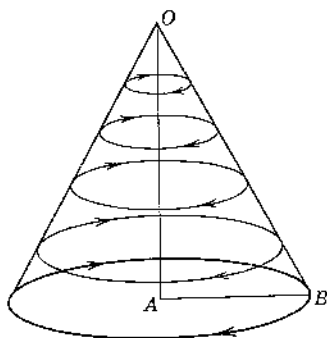


FIG. 15

will be considered in Art. 5. Solids meeting this condition may be designated as solids with a *slice formula*, for reasons that will become apparent after the discussions in Arts. 4 and 5.

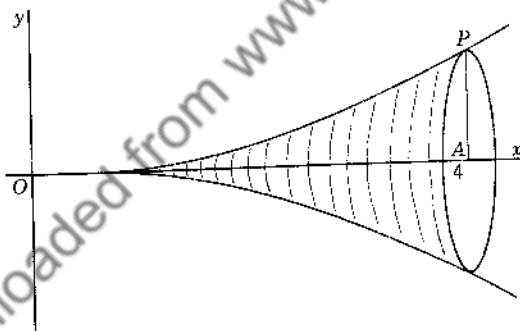


FIG. 16

4. Volumes of Revolution. The first figure of revolution we shall study is the cone generated by revolving the area OAB in Fig. 17 about the x -axis. Let the equation of the line OB be $y = \frac{1}{3}x$, and let the altitude of the cone be a . The radius of the base is then $\frac{1}{3}a$. A familiar formula from solid geometry gives the volume of a cone: $V = \frac{1}{3}\pi r^2 h$, where r is the radius of the base and h is the altitude. By this formula, the volume of the given cone is $V = \frac{1}{3}\pi(\frac{1}{3}a)^2 \cdot a = \frac{\pi a^3}{27}$.

Here, no new method of finding a volume is needed. The formula, however, was derived by a process essentially equivalent to the general method we are about to discuss.

If we knew no elementary formula for the volume of a cone the solution of the area problem might suggest the following attack. An approximation to the volume is given by the sum of the volumes of the cylindrical slices drawn in Fig. 17. Such an approximating sum is formed by subdividing the interval $0 \leq x \leq a$ into several subintervals; then, with each subinterval as an altitude, inscribing a cylinder in the cone; and, finally, adding together the cylindrical volumes thus formed. We might,

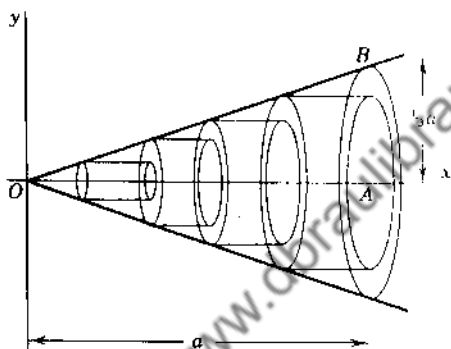


FIG. 17

equally well, circumscribe the approximating cylinders. If we let n be the number of subintervals, $\underline{S}_n$ the sum of the volumes of the inscribed cylinders, and $\overline{S}_n$ the sum of the volumes of the circumscribed cylinders, we would be intuitively satisfied that the volume we seek lies between the two sums. That is,

$$\underline{S}_n < V < \overline{S}_n. \quad (4.1)$$

If we consider approximating sums, $\underline{S}_n$ and $\overline{S}_n$, formed with larger and larger values of n , the lengths of all of the subintervals tending to 0 as n tends to infinity, it is intuitively apparent that the difference between $\overline{S}_n$ and $\underline{S}_n$ becomes less and less and approaches 0. On account of inequality (4.1), which is also intuitively apparent for all values of n , we are led to define the volume V as the common limit $\underline{S}_n$ and $\overline{S}_n$. This definition is completely analogous to the definition of area given in Chapter 7.

The fact that $\underline{S}_n$ and $\overline{S}_n$ approach a common limit, just as in the case of the area problem, follows from our earlier discussions and needs no

new proof.* In fact, this common limit is a definite integral. Let us set up the sum that leads to this definite integral. For this purpose we consider a sum of approximating cylinders, each with its radius not less than the least ordinate of the line $y = \frac{1}{3}x$ in the subinterval and not greater than the greatest ordinate in this subinterval. If we denote the length of the subinterval by $\Delta_i x$, we then have to express the volume of the cylinder of altitude $\Delta_i x$ and radius $y_i = \frac{1}{3}x_i$, where x_i is a value of x in the subinterval. From the formula for the volume of a cylinder,† $V = \pi r^2 h$, we have for the volume of one of our approximating cylinders,

$$\pi\left(\frac{1}{3}x_i\right)^2 \Delta_i x. \quad (4.2)$$

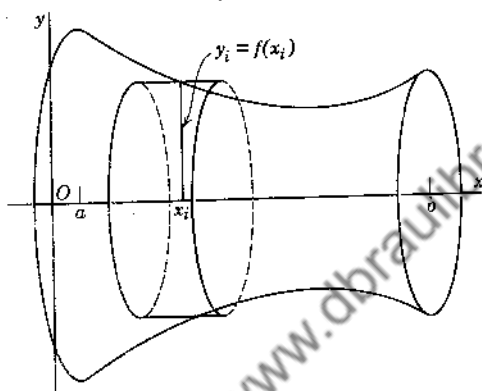


FIG. 18

Expression 4.2 represents the volume of an elemental slice of the volume of the approximating solid. Since the volume of each such slice is given by a function of one variable x , multiplied by the thickness $\Delta_i x$, this solid has a slice formula and the condition laid down in Art. 3 is satisfied.

The sum that approximates the volume of the cone is now $\sum_{i=1}^n \pi\left(\frac{1}{3}x_i\right)^2 \Delta_i x$. This sum has the form $\sum_{i=1}^n f(x_i) \Delta_i x$, where $f(x) = \pi\left(\frac{1}{3}x\right)^2$. Its limit, as n tends to infinity, is therefore the integral

$$\int_0^a \pi\left(\frac{1}{3}x\right)^2 dx = \int_0^a \frac{\pi}{9} x^2 dx.$$

Evaluating this integral, we obtain $\pi a^3/27$, the same result as that given by the familiar formula.

* Of course this was not proved in this book. The proof referred to here is that of the existence of the definite integral of a continuous function. See Art. 9, Chapter 7.

† This formula, like that for the volume of a cone, is derived in solid geometry by a limiting process, essentially the same as that which leads to the formula for the area of a circle.

We now generalize this discussion. Let an area be bounded by the graph of the equation $y = f(x)$, the x -axis, and two ordinates $x = a$ and $x = b$. We wish to find the volume of revolution generated by revolving this area about the x -axis (Fig. 18). We define this volume by the definite integral $\int_a^b \pi [f(x)]^2 dx$. Since the radius of an elemental cylindrical slice of this volume is naturally expressed as the length of the ordinate y , the expression for the volume is often written $\int_a^b \pi y^2 dx$, with the understanding that y is to be replaced by $f(x)$ before integration. Thus

$$V = \int_a^b \pi y^2 dx = \int_a^b \pi [f(x)]^2 dx = \pi \int_a^b [f(x)]^2 dx. \quad (4.3)$$

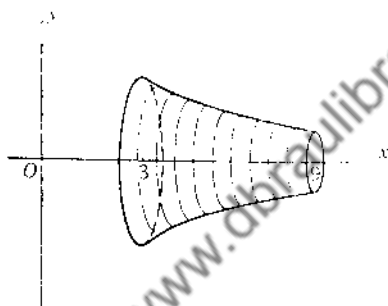


FIG. 19

EXAMPLE 1. Find the volume generated by revolving about the x -axis the area bounded by the parabola $y = x^2/16$, the x -axis, and the line $x = 4$ (Fig. 16).

SOLUTION. We have, from (4.3),

$$\begin{aligned} V &= \pi \int_0^4 (x^2/16)^2 dx = (\pi/256) \int_0^4 x^4 dx \\ &= (\pi/256) \left[x^5/5 \right]_0^4 = 4\pi/5. \end{aligned}$$

EXAMPLE 2. Find the volume generated by revolving about the x -axis the area bounded by the hyperbola $xy = 9$, the x -axis, and the ordinates $x = 3$ and $x = 9$ (Fig. 19).

SOLUTION. From formula (4.3)

$$\begin{aligned} V &= \pi \int_3^9 (9/x)^2 dx = 81\pi \int_3^9 1/x^2 dx \\ &= 81\pi \left[-1/x \right]_3^9 = 81\pi \left(-\frac{1}{9} + \frac{1}{3} \right) = 18\pi. \end{aligned}$$

EXAMPLE 3. Find the volume generated by revolving about the y -axis the area bounded by the curve $y = x^2$, the y -axis, and the line $y = 4$ (Fig. 20).

SOLUTION. The cylindrical slices by which this volume is approximated have the radii x_i , instead of y_i , where $x_i^2 = y_i$, or $x_i = \sqrt{y_i}$. We can obtain the volume by replacing x by y in formula 4.3. Thus,

$$\begin{aligned} V &= \pi \int_0^4 x^2 dy = \pi \int_0^4 y dy \\ &= \pi \left[\frac{y^2}{2} \right]_0^4 = 8\pi. \end{aligned}$$

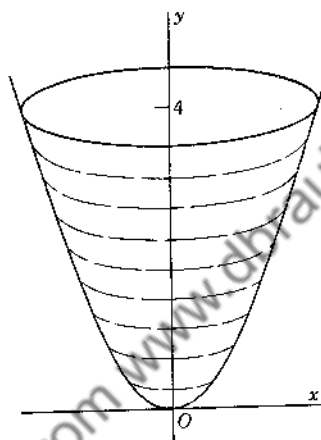


FIG. 20

EXERCISES

Find the volume of revolution generated by revolving about the x -axis the area bounded by the graph of $y = f(x)$, the x -axis, and the ordinates given, in each of Exercises 1-10.

1. $f(x) = x^2$,
2. $f(x) = 1/x^2$,
3. $f(x) = \frac{1}{2}\sqrt{x}$,
4. $f(x) = \frac{1}{4}(x+1)^2$,
5. $f(x) = \sqrt{x+4}$,
6. $f(x) = 9x - 2x^2$,
7. $f(x) = x + 1/x$,
8. $f(x) = -x^2$,
9. $f(x) = x^3 + 5x^2 - 6x$,
10. $f(x) = x^3 + 5x^2 - 6x$,

- $x = 1$.
- $x = 1, x = 2$.
- $x = 6$.
- $x = 0$.
- $x = 0$.
- $x = 3$.
- $x = 1, x = 4$.
- $x = 2$.
- $x = -6, x = 0$.
- $x = -6, x = 1$.

11. Find the volume generated by revolving about the x -axis the area bounded by the parabola $y = 9 - x^2$ and the straight line $y = 8$. [Hint: Find two volumes and subtract.]

12. Find the volume generated by revolving the area bounded by the parabola $y^2 = x$ and the straight line $y = x$, (a) about the x -axis; (b) about the y -axis.

13. Find the volume generated by revolving about the x -axis the area bounded by the parabola $y^2 = x$, the y -axis, and the straight line $y = 2$.

14. Find the volume generated by revolving about the x -axis the area bounded by the parabola $y = 7 - x^2$ and the hyperbola $xy = 6$.

15. Find the volume in Exercise 14 if the area is revolved about the y -axis.

16. Find the volume of a segment of a sphere if the radius of the sphere is a and the bases of the segment are at distances m and n from the center and are on the same side. [Hint: Revolve the semicircle $y = \sqrt{a^2 - x^2}$ about the diameter.]

17. Use the answer to Exercise 16 to obtain (a) a formula for the volume of a segment of one base with its thickness equal to h ; (b) a formula for the volume of a segment of thickness h with one base through the center of the sphere.

18. Find the volume of the ellipsoid of revolution generated by revolving the ellipse $x^2/a^2 + y^2/b^2 = 1$ about its major axis. (This solid is called a *prolate spheroid*.)

19. Find the volume of the ellipsoid generated by revolving the ellipse of Exercise 18 about its minor axis. (This solid is called an *oblate spheroid*.)

20. The area bounded by the parabolas $x^2 = y$ and $y^2 = x^2$ and the coordinate axes is revolved about one of these axes. Find the volume generated.

5. Other Solids Having Slice Formulas. Let us try to calculate the volume of the figure shown in Fig. 14. We think of this volume as being approximated by the sum of slices, each of which is a thin plate with equal square bases and narrow rectangular sides. One such slice is

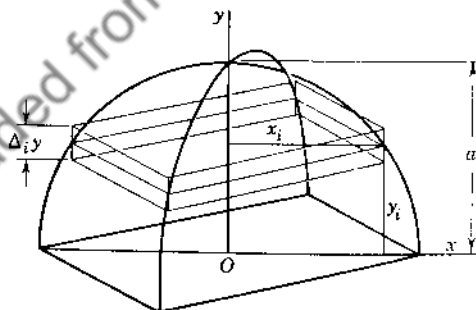


FIG. 21

pictured in Fig. 21. From the discussion in Art. 3, we know that the area of the base of this elemental slice is equal to $2(a^2 - y_i^2)$ where y_i is the height of the slice above the x -axis. This height can be measured to the bottom of the slice, the top, or any point between. If we call the thickness of the slice Δy , the slice formula is

$$\text{Volume of the slice} = 2(a^2 - y_i^2)\Delta y.$$

Summing slices corresponding to n subintervals of the y -axis between $y = 0$ and $y = a$, we obtain

$$\sum_{i=1}^n 2(a^2 - y_i^2) \Delta y.$$

The limit of this sum, as n tends to infinity, is $\int_0^a 2(a^2 - y^2) dy$. As for volumes of revolution, we define the volume by this integration. Thus,

$$V = \int_0^a 2(a^2 - y^2) dy = (2a^2y - 2y^3/3) = 4a^3/3.$$

As a second example, let us calculate the volume of the solid in Fig. 22. The base of this solid is the figure bounded by the parabola $y^2 = x$ and

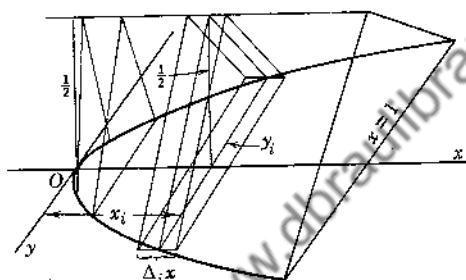


FIG. 22

the line $x = 1$. Every cross section perpendicular to the x -axis is an isosceles triangle with altitude $\frac{1}{2}$. Now, consider any particular cross section perpendicular to the x -axis, at a distance x_i from the origin. Think of an elemental slice of the solid as a prism having this triangular cross section as base, and having the thickness Δx . The length of the base of the triangular cross section is $2y_i$, where $y_i = \sqrt{x_i}$, and the altitude is $\frac{1}{2}$. Therefore, the volume of the slice is

$$\begin{aligned} \text{Volume of the slice} &= \frac{1}{2}(2y_i)\frac{1}{2}\Delta x \\ &= \frac{1}{2}y_i\Delta x = \frac{1}{2}\sqrt{x_i}\Delta x. \end{aligned}$$

This is the slice formula. Summing the volumes of slices corresponding to n subintervals of the x -axis from $x = 0$ to $x = 1$, we obtain for the

approximate volume $\sum_{i=1}^n \frac{1}{2}\sqrt{x_i}\Delta x$. The limit of this sum, as n tends to

infinity, is the integral $\int_0^1 \frac{1}{2}\sqrt{x} dx$, and this is the volume of the solid, by definition. Hence,

$$V = \int_0^1 \frac{1}{2}\sqrt{x} dx = \left[\frac{1}{3}x^{3/2} \right]_0^1 = \frac{1}{3}.$$

The steps in calculating a volume for which there is a slice formula of the kind described are (1) determine a plane, such that the area of every cross section of the solid parallel to this plane can be expressed as a function of its distance, say x , from the plane (be sure that this formula gives the correct value of the area for all cross sections); (2) write down the slice formula (if s_i denotes a value of x in a typical slice, this is the product of the function of s_i that is equal to the area of the cross section multiplied by Δs_i ; that is, $f(s_i)\Delta s_i$); (3) express the limit of the sum of such slices as the integral $\int_a^b f(x) dx$, in which a and b are, respectively, the lower and upper limits of the values of x that correspond to points of the solid; and (4) evaluate this integral.

Frequently, all the cross sections used in this method are similar figures. This is true of figures of revolution, which have circular cross sections. It was true of the first example in the present article, in which the cross sections were squares. That it need not be so is illustrated by the second example.

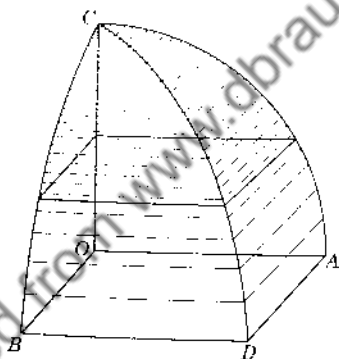


FIG. 23

EXERCISES

1. The base of a solid is the figure bounded by the parabola $y^2 = x$ and the line $x = 1$. Every cross section perpendicular to the x -axis is a square. Find the volume.
2. Find the volume in Exercise 1 if the cross sections perpendicular to the x -axis are equilateral triangles.
3. The base of a solid is a circle with radius 2. Every cross section perpendicular to a given diameter is an isosceles triangle with its altitude equal to its base. Find the volume.
4. Find the volume of the solid in Fig. 23, given that OAC and OBC are quadrants of circles of radius 4 with their planes at right angles, and that every section parallel to the base $OBDA$ is a square. Hence, find the volume common to two circular cylinders, each having its radius equal to 4 units, and with their axes intersecting at right angles.
5. Find the volume of the solid such that every cross section perpendicular to the x -axis is a circle of radius 2 with its center on the semicircle $y = \sqrt{4 - x^2}$. Compare this volume with that of the cylinder of radius 2 and altitude equal to 4.
6. Find the volume of the solid such that every cross section perpendicular to the

x -axis is a circle with its center on the semicircle $y = \sqrt{4 - x^2}$ and its circumference intersecting the x -axis.

7. Show that the volume of any circular cone is equal to $\pi r^2 h/3$, in which r is the radius of the base and h is the altitude, whether or not the axis of the cone is perpendicular to the base. Show that the horn-shaped solid in Fig. 24 also has this volume, if every

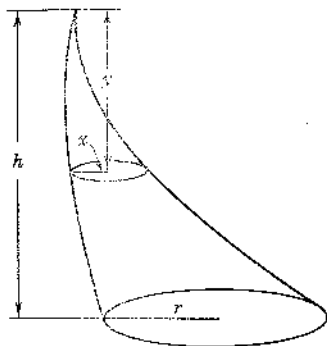


FIG. 24

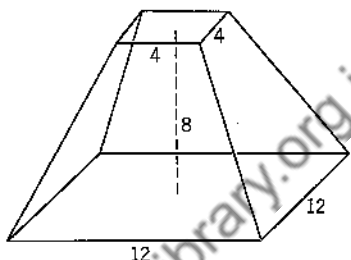


FIG. 25

horizontal cross section is a circle with its radius proportional to the length of the perpendicular from the vertex.

8. Find the volume of the frustum of a square pyramid with the two bases 12 in. and 4 in. on a side, respectively, and the altitude 8 in., as in Fig. 25.

9. Find the volume of the wedge $ABCD$ cut from a circular cylinder of radius r , if one of the two planes forming the wedge cuts perpendicularly across the cylinder

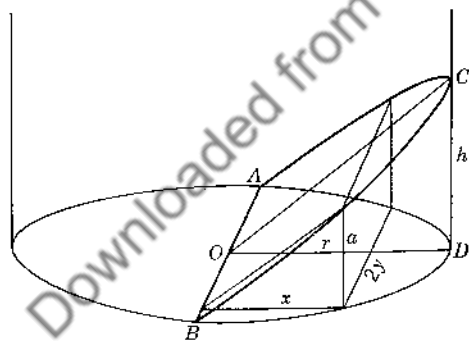


FIG. 26

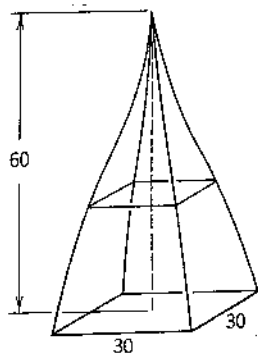


FIG. 27

and the other meets it along a diameter at an angle $\tan^{-1} h/r$ (Fig. 26). [Hint: The integral giving this volume may be set up equally well in two different ways: (1) by using rectangular slices perpendicular to OD , and (2) by using right triangular slices perpendicular to AB . In either method, use the equation of the circular base, $x^2 + y^2 = r^2$, and the proportion $a/x = h/r$.]

10. A tower is 60 ft high and 30 ft square at the base, and every cross section parallel to the base is square (Fig. 27). The side, s feet, of any of these cross sections is given by the formula $s = (15x^2 - x^3/6)/60$, where x is the distance in feet from the peak. Find the volume of the tower.

6. Work. The problem of finding the work done by a force in moving a body a distance along a line was described in Art. 10, Chapter 1, and the mathematical aspect of this problem was shown to be identical with that of the area problem. Now that the area problem has been solved and, indeed, has been generalized through the formulation of the definite integral, we shall consider the work problem anew. We shall not rely upon the identification of these two problems as discussed in Chapter 1. Rather, an integral will be constructed, by direct reference to the physical quantities involved, to represent what is intuitively acceptable as the amount of work. We consider a concrete problem.

According to Hooke's law, the force necessary to stretch a spring from its natural length s_0 to length s is proportional to the amount the spring is stretched. That is,

$$F = k(s - s_0). \quad (6.1)$$

Thus, the force varies with s . If a definite force could act instantaneously, the spring would instantly stretch to a length depending on the force, in accordance with this formula. But Hooke's law implies that the force varies during the process of stretching, so that at the beginning it is 0 and gradually increases to its maximum which, in turn, determines the total amount the spring is stretched, in accordance with (6.1).

Suppose a certain spring that is 10 in. long when no force is applied is stretched to the length of 12 in. by a force that reaches 20 lb. Let us try to find the amount of work done by the force in stretching the spring this 2 in. Formula 6.1 is a general formula that applies to any spring. We first determine the precise formula for this spring, giving the amount of the force acting at any point, i.e., at any length of the spring. For this purpose we put $s_0 = 10$, so that $F = k(s - 10)$, and observe that when $s = 12$ $F = 20$. Hence, $20 = k(12 - 10)$, and $k = 10$. For this spring, then,

$$F = 10(s - 10). \quad (6.2)$$

The value of k is called the *spring constant* belonging to this particular spring.

Now, if the force were constantly 20 lb, the work would be $20 \cdot 2 = 40$ in.-lb, since the definition given in physics for the work done by a constant force F in moving a body s units of distance is $F \cdot s$. But the force is not constant, for it starts at 0 and reaches 20 only at the end of the interval. Hence, $20 \cdot 2$ appears to be a poor approximation to the work actually done. We can no more multiply a varying force by the distance, to find the actual work, than we could multiply a varying height of a curve by the length of a segment of the x -axis to find the area under the curve. Hence, the definition $W = F \cdot s$ does not apply to a

varying force. We therefore resort to an approximation in the formulation of which the definition does apply. We imagine the interval of 2 in. to be subdivided into n short subintervals of lengths $\Delta_1 s$, $\Delta_2 s$, etc., and in each subinterval $\Delta_i s$ we imagine the variable force to be replaced by a constant force, F_i . If $\underline{F}_i$ and $\overline{F}_i$ denote the minimum and maximum, respectively, of the variable F in the subinterval, we shall make F_i meet the requirement $\underline{F}_i < F_i < \overline{F}_i$. The work done by the substitute force F_i in stretching the spring $\Delta_i s$ units is $F_i \cdot \Delta_i s$, and the work done by all substitute forces in stretching the spring 2 in. is

$$\sum_{i=1}^n F_i \Delta_i s. \quad (6.3)$$

The use of these substitute forces implies, of course, a substitute formula that only approximates Hooke's law. The approximation is, in general, improved by increasing n , provided that every $\Delta_i s$ decreases as n increases.

As $n \rightarrow \infty$ the sum (6.3) tends to the integral $\int_{10}^{12} F ds$. The work done in stretching the spring to the length 12 in. is defined by this integral:

$$W = \int_{10}^{12} F ds.$$

Since $F = 10(s - 10)$, by (6.2),

$$W = \int_{10}^{12} 10(s - 10) ds = 20 \text{ (in.-lb.)}$$

In the same way, in general, the work done by a variable force $F = f(s)$, acting through a distance from a to b on a straight line is defined to be

$$W = \int_a^b f(s) ds. \quad (6.4)$$

We shall consider some additional examples.

EXAMPLE 1. Coulomb's law states that the force of attraction between two unlike electrical charges varies inversely as the square of the distance between them. If the force is 3 dynes when the charges are 4 cm apart, find the work done in moving them to a distance of 12 cm apart.

SOLUTION. We have $F = k/x^2$, where F is in dynes and x is the number of cm that separate the charges. When $x = 4$, $F = 3$. Therefore,

$$3 = k/16, \quad k = 48, \quad \text{and} \quad F = 48/x^2.$$

Using Eq. 6.4 to find the work done in increasing the distance between the charges from 4 cm to 12 cm,

$$W = \int_4^{12} 48/x^2 \, dx = \left[-48/x \right]_4^{12} = -4 + 12 = 8.$$

The amount of work is therefore 8 dyne-cm.

EXAMPLE 2. Under certain conditions that are usually approximated in a gas engine, the product of the number of units of pressure per square unit of area and the number of units of volume of gas are related by the equation

$$pv^{1.4} = k = a \text{ constant.}$$

If the gas is confined in a cylinder with a piston at one end, the pressure of the gas causes the piston to move so as to increase the volume, thus

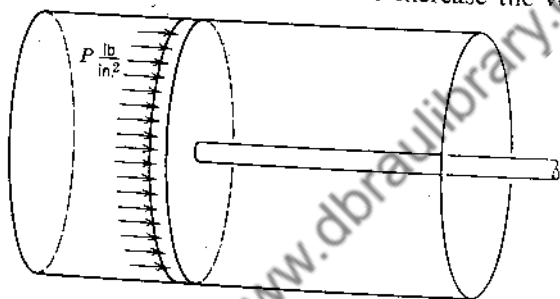


FIG. 28

lowering the pressure (Fig. 28). Suppose the diameter of the cylinder is 12 in. and the pressure is 100 lb per sq in. when the piston is 10 in. from the inner end of the cylinder. Find the work done by the force due to the pressure of the gas in moving the piston 12 in., i.e., from 10 in. to 22 in. from the inner end.

SOLUTION. Since $pv^{1.4} = k$, we first find k from the fact that $p = 100$ when $v =$ the volume of the cylinder 12 in. in diameter and 10 in. long. To simplify notations, we denote the area of the piston, which is 36π sq in., by A . We have $p = 100$ when $v = 10A$, so that $k = 100(10A)^{1.4}$. Letting x represent the variable length of the cylindrical chamber, we have $v = Ax$, and, hence,

$$p = \frac{k}{(Ax)^{1.4}} = \frac{100(10A)^{1.4}}{(Ax)^{1.4}} = \frac{100(10)^{1.4}}{x^{1.4}}.$$

Since p denotes the pressure per square inch, the force acting on the piston is the product pA ; i.e.,

$$F = \frac{100(10)^{1.4}A}{x^{1.4}}.$$

Therefore,

$$\begin{aligned} W &= \int_{10}^{22} \frac{100(10)^{1.4} A}{x^{1.4}} dx \\ &= 100(10)^{1.4} A \left[-1/.4x^{.4} \right]_{10}^{22} \\ &= 100(10)^{1.4} A (10/4) \left[-1/x^{.4} \right]_{10}^{22} \\ &= 250(10)^{1.4} A [1/(10)^{.4} - 1/(22)^{.4}]. \end{aligned}$$

By logarithms,

$$10^{1.4} = 25.12 \quad \text{and} \quad 1/(10)^{.4} - 1/(22)^{.4} = .1077.$$

Hence,

$$W = 250(25.12)(.1077)A = 676.3A.$$

Putting $A = 36\pi$ in this result,

$$W = 76,490 \text{ in.-lb} \quad \text{or} \quad 6374.2 \text{ ft.-lb.}$$

EXERCISES

1. A mass is moved along a straight line by a force F lb, with $F = 1 + \sqrt{s}$, s being the distance in feet from some fixed point of the line. Find the work done by this force in moving the mass from $s = 1$ to $s = 9$.

2. Find the work done in stretching the spring in the example discussed above 2 in. more, i.e., from length 12 in. to length 14 in.

3. A force of 60 lb stretches a spring 16 in. long to a length of 20 in. Find the work done in stretching the spring 1 in. more.

4. The force of gravitation between two bodies is

$$F = km_1m_2/s^2,$$

in which m_1 and m_2 are the masses of the bodies, s is the distance between their centers, and k is a constant. The force of gravitational attraction between the earth and a 1-lb mass on its surface is 1. Find the work done against gravity in moving a 1-lb mass from the surface of the earth to a point 1000 miles above the surface. Assume the earth's radius to be 4000 miles.

5. How much work is done against gravity in propelling a rocket weighing T tons at the earth's surface to a height of 100 miles above the earth?

6. The force of repulsion between two like electrical charges is k/s^2 , where s is the distance between them and k is a constant. If the force is 6 dynes when the charges are 10 cm apart, how much work is done by the force of repulsion in separating the charges to a distance of 40 cm?

7. A steam engine has a cylinder 24 in. long and 12 in. in diameter. At the beginning of the forward stroke the piston is at the backward end of the cylinder. During the first third of the stroke steam with a boiler pressure of 100 lb per sq in. is admitted into the cylinder. The opening from the boiler into the cylinder is then closed and the steam expands in such a way that $pv^{1.4} = k$. Find the work done on the piston during the forward stroke. [Hint: During the first third of the stroke, or until the piston

reaches the position $ABCD$ in Fig. 29, the force is constant, since the pressure is constant. Evaluate k at this point.]

8. A quantity of gas in a cylinder with an initial volume of 4 cu ft and pressure of 10 lb per sq in. is compressed by means of a piston until the volume is 1 cu ft. If the formula relating volume and pressure is $pv^{1.1} = k$, what is the final pressure and how much work is done? [Hint: The diameter of the cylinder is immaterial.]

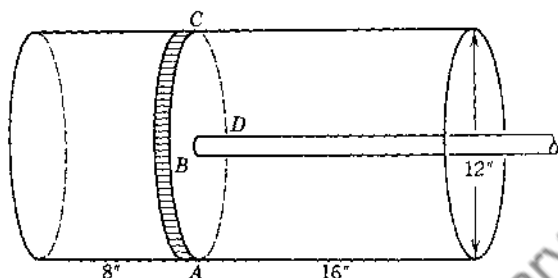


FIG. 29

7. Liquid Pressure on a Vertical Wall. The amount of the force due to the pressure of a liquid on the bottom of a container is easily calculated. It is simply the weight of all the liquid resting on it. The total force due to the pressure on the vertical sides of a container is less easily found. It is a matter of common experience that the pressure does exert a force

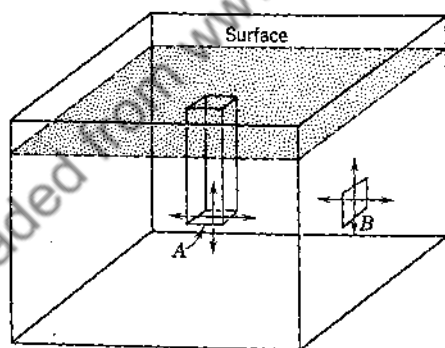


FIG. 30

on the vertical side, and it is also experimentally obvious that the force on 1 sq in. near the surface of the liquid is less than that on the same area at a deeper level. Two physical laws are illustrated by these observations: (1) at any point in a motionless liquid the pressure exerts the same force in all directions; and (2) at any depth, the force due to pressure on a unit area horizontally placed is equal to the weight of the vertical column of liquid directly above this area. At the point A in the tank pictured in Fig. 30 the pressure exerts equal forces in all directions. Because of the

balancing forces, the particle of liquid at A is in equilibrium and does not move. At B , the liquid exerts the same force against the wall of the tank that it exerts downward and in all other directions. But although it is easy to compute the downward force on the square base of the column at A by measuring the weight of the column above it, there is no column resting on the vertical square centered at B .

We might reason that since the force is the same in all directions, we can just as well think of the square as being horizontal. This might serve if we could decide at what depth to place it. If we make it horizontal at the depth of its lower edge we are assuming that the force at every point on it has the value that it has at its greatest depth. This is clearly wrong. It would be equally wrong to think of it as horizontal at the depth of the

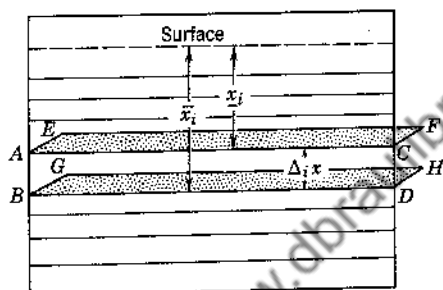


FIG. 31

upper edge. The first assumption would make the force too great, and the second would make it too small. Thus, some intermediate depth ought to give us the correct value. What that depth is we cannot know in advance. If this could be determined we could sum up the forces on all units of area of the vertical wall to obtain the total force.

Just as with the tangent problem and the area problem, although we do not have the exact answer to the question of the total force on the vertical side of the tank, we can obviously approximate it. Let us divide the surface of the side of the tank into strips, as illustrated in Fig. 31. The force of the pressure on a strip $ABCD$ is more than it would exert on a strip $ACEF$ exactly like this one but placed horizontally at the level of AC . At the same time it is less than it would exert on a strip $BDGH$ exactly like this one but placed horizontally at the level of BD . If we imagine every strip rotated so as to lie horizontally in the level of its upper boundary and form the sum of the quantities representing the forces due to pressure on the tops of these strips, we shall have a sum $\bar{S}_n$ that is less than the total force of the pressure on the wall. Similarly, we can form a sum $\bar{S}_n$ that is greater than this force, by thinking of each

strip rotated so as to lie horizontally in the level of its lower boundary. If P represents the actual force against the wall, we then have

$$\underline{S}_n < P < \overline{S}_n \quad (7.1)$$

Inequality (7.1) has a familiar appearance, and we should expect it to lead us to the formulation of the value of P as a definite integral. In order to do this we must express $\underline{S}_n$ and $\overline{S}_n$ in terms of the numerical quantities of the problem. Now, the weight W of the column (it is really a slab) of liquid above the rectangle $ACEF$ is equal to the volume of the column multiplied by the density δ , or weight per unit volume. Since the volume is equal to the area of the rectangle multiplied by its depth below the surface,

$$W = \delta l w x_i,$$

where l , w , and x_i , represent the length, width, and depth of the rectangle, respectively. Since each rectangle used in forming $\underline{S}_n$ was obtained by rotating the rectangle $ABCD$ about its upper boundary, $w = \Delta_i x$, and $x_i = \underline{x}_i$, the least depth of any point of the original rectangle. Thus,

$$W = \delta l \underline{x}_i \Delta_i x.$$

In the present example, l is a constant for all rectangular strips, and we assume the density to be constant. The sum $\underline{S}_n$, corresponding to the least depths of all of the strips, $\underline{x}_i$, is

$$\underline{S}_n = \sum_{i=1}^n \delta l \underline{x}_i \Delta_i x.$$

Similarly,

$$\overline{S}_n = \sum_{i=1}^n \delta l \overline{x}_i \Delta_i x.$$

Here, x_i is the greatest depth of any point in the corresponding strip. Inequality (7.1) can now be written

$$\underline{S}_n = \sum_{i=1}^n \delta l \underline{x}_i \Delta_i x < P < \sum_{i=1}^n \delta l \overline{x}_i \Delta_i x = \overline{S}_n. \quad (7.2)$$

The sums in inequality (7.2) are exactly the kind of sums that we have considered before in defining the definite integral. As n tends to infinity and each width $\Delta_i x$ tends to 0, $\underline{S}_n$ and $\overline{S}_n$ tend to the same limit, which must therefore be the limit approached by any sum, $\sum_{i=1}^n \delta l x_i \Delta_i x$, in which x_i represents any depth between $\underline{x}_i$ and $\overline{x}_i$. Since

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \delta l x_i \Delta_i x = \int_0^h \delta l x \, dx,$$

where h is the depth of the bottom of the tank, the total force of the pressure on the vertical wall is given by this integral.

If $l = 10$ (feet), $h = 5$ (feet), and $\delta = 62.5$ (lb per cu ft), which is the density of water,

$$\begin{aligned} P &= \int_0^5 (62.5) \cdot (10) \cdot x \, dx \\ &= 625(25/2) = 7812.5. \end{aligned}$$

Hence, the total force due to pressure of the liquid on the side of the tank is 7812.5 lb.

An added complication arises when the side of the container is not rectangular. Suppose, for example, that it is semicircular (Fig. 32). As before, we divide this area into strips. These are no longer rectangular strips. But we can approximate them by rectangles, each with its length l_i lying between the least and greatest length of the strip. We now set

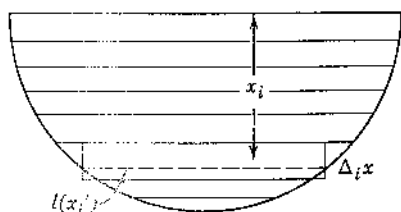


FIG. 32

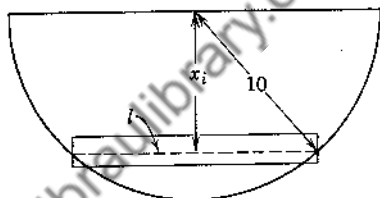


FIG. 33

up the approximating sum $\sum_{i=1}^n \delta l(x'_i) x_i \Delta_i x$, where $l(x'_i)$ is the length of the side of the container at depth x'_i within the strip, and x_i is also a depth within the strip, not necessarily the same as x'_i . This sum has a form to which reference was made near the end of Art. 9, Chapter 7, and its limit is the integral $\int_0^h \delta \cdot l(x) \cdot x \, dx$.

For the example in question, suppose that the radius of the semicircular side is 10 ft, that the liquid is water, and that the tank is full. From Fig. 33 it is clear that $l(x) = 2\sqrt{100 - x^2}$. Since $\delta = 62.5$, we have for the total force,

$$\begin{aligned} P &= \int_0^{10} (62.5)(2\sqrt{100 - x^2}) x \, dx. \\ &= 62.5 \left[-\frac{2}{3}(100 - x^2)^{\frac{3}{2}} \right]_0^{10} \\ &= -\frac{125}{3} (0 - (100)^{\frac{3}{2}}) \\ &= \frac{125}{3} \cdot 1000 = 41,667, \text{ approximately.} \end{aligned}$$

Hence the force on the side of the tank is approximately 41,667 lb.

EXERCISES

1. A rectangular gate in a vertical dam is 12 ft wide and 6 ft high, and its top is 9 ft below the level of the water. Find the total force due to the pressure of the water on the gate. (Density of water = 62.5 lb per cu ft.) [Hint: Let x = depth below the water level. Then

$$P = \int_0^{15} 62.5 \cdot 12 \cdot x \, dx.]$$

2. If the water rises 6 ft what will the force in Exercise 1 then be?

3. The vertical end of a horizontal tank is a circle 8 ft in diameter. (a) Find the force on this end due to the pressure of water, if the tank is one-half full. (b) Find the force if the tank is one-half full of oil weighing 56 lb per cu ft.

4. The vertical end of a horizontal oil tank is an ellipse, of which the major axis is 10 ft long and placed horizontally, and the minor axis is 6 ft long. Find the force on this end due to the pressure of oil in the tank if the tank is one-half full of oil weighing 56 lb per cu ft.

5. A trapezoidal gate in a vertical dam has its longer base 12 ft long and 5 ft below the surface of the water, and its shorter base 8 ft long and 11 ft below the surface. Calculate the force on the gate due to the water pressure. [N.B. It makes no difference whether or not the trapezoid is isosceles.]

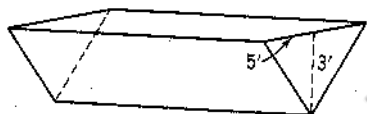


FIG. 34

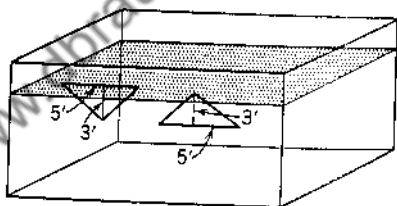


FIG. 35

6. The vertical end of a water trough is an isosceles triangle 5 ft across the top and 3 ft deep (Fig. 34). Calculate the force on one end due to the water pressure if the trough is full of water.

7. A board in the shape of an isosceles triangle, having its base 5 ft long and its altitude 3 ft long, is submerged vertically, so that its base is horizontal and lies in the surface of the water. (a) Find the force on one side of the board. (b) Find the force on one side of the board if it is turned upside down (Fig. 35).

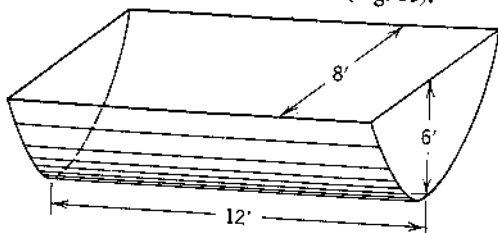


FIG. 36

8. The vat pictured in Fig. 36 has parabolic ends. Find the force on one end due to the pressure of liquid when the vat is full of a liquid weighing 60 lb per cu ft.

Trigonometric Functions

1. Introduction. The preceding chapters constitute an introduction to differential and integral calculus. In order that all possible emphasis might be put on the ideas and processes that are characteristic of calculus, difficulties that are not inherent in the calculus operations themselves were avoided as much as possible. It is for this reason that the discussions up to this point have been confined to relatively simple algebraic functions. We must now extend the field of study to include some more advanced functions. In applying the limiting operations to these functions some new difficulties will be encountered. The understanding of the limiting operations as applied to simpler functions in the earlier chapters, however, will smooth the way so that we can give most of our attention to the special problems that these new functions place before us.

The new functions that we shall study are by no means unfamiliar, having been introduced in a concrete way in elementary work. In the present chapter we shall extend the calculus operations to the trigonometric functions and their inverses, whereas Chapter 11 will be devoted to the corresponding extension to logarithmic functions and their inverses, exponential functions. These functions, together with the hyperbolic functions, which will be taken up at a later point, are called the elementary transcendental functions.

The applications of transcendental functions are almost as frequent as those of algebraic functions. For example, the trigonometric functions are especially important in the study of oscillatory motion, and the logarithmic and exponential functions have many applications in studies of growth and decay.

2. Radian Measure. Since angles and their measures are at the basis of our definitions of the trigonometric functions, we shall review some essential ideas about units by which angles are measured. The most natural unit of angular measure is a complete revolution. The familiar degree is defined simply as $1/360$ of a revolution. A minute is then defined as $1/60$ of a degree, and a second as $1/60$ of a minute.

A less familiar unit, which is fundamental in the work of the present chapter, is the *radian*. Imagine an angle with vertex at the center of a circle and sides cutting off an arc of the circumference equal in length to the radius (Fig. 1). This angle is called a radian. If a central angle of θ radians intercepts an arc of length s , the number of radians in the angle is s/r ; i.e., $\theta = s/r$, or $s = \theta r$.

Relations connecting the number of degrees in an angle with the number of radians in it are easily worked out. Thus, since the circumference $= 2\pi$ times the radius, the angle of $360^\circ =$ the angle of 2π rad;

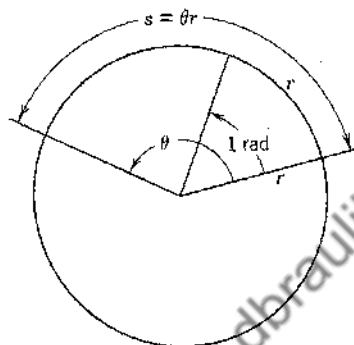


FIG. 1

the angle of $1^\circ =$ the angle of $2\pi/360$, or $\pi/180$, rad; and the angle of $(180/\pi)^\circ =$ the angle of 1 rad. Several often-used angles are shown in both degree measure and radian measure in the following table.

Radians	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	$2\pi/3$	$3\pi/4$	$5\pi/6$	π	1	$\pi/180$
Degrees	0	30	45	60	90	120	135	150	180	$180/\pi$	1

EXERCISES

1. In a circle of radius 2 units, what is the length of arc intercepted by a central angle of (a) 1 rad? (b) 3 rad? (c) $\frac{3}{4}$ rad? (d) 7 rad?
2. A wheel makes 50 revolutions per min. What is its speed of rotation in radians per second? in degrees per second?
3. A wheel of radius 4 in. revolves at 30 rad per sec. How fast is a point on the rim moving?
4. A wheel of radius 16 in. rolls along the ground at 30 mph. What is its speed of rotation in radians per second?
5. The earth rotates through 1 revolution in 24 hr. (a) How fast is this in radians per second? (b) If the radius is 4000 miles, how fast does a point on the equator move as a result of this rotation? (c) How fast does a point at latitude 60° move as a result of this rotation?

3. The Functional Relation $y = \sin x$. The definitions of the trigonometric functions, or trigonometric ratios, as given in elementary trigonometry, are dependent on geometrical concepts. Indeed, the geometrical notion of an angle is at the very core of trigonometry, and the trigonometric functions are thought of as functions of an angle. For example, the "sine" is meaningless except as the "sine of an angle."

A general definition of the sine is made by placing the angle A (of any size, positive or negative) on a rectangular coordinate system, in "standard position," as illustrated in Fig. 2. After this is done, some point on the

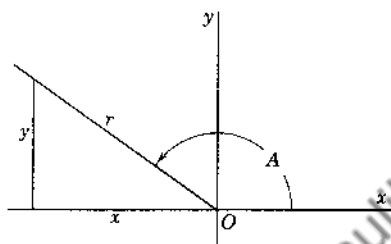


FIG. 2

terminal side is chosen (any point will do except the vertex), and the coordinates x and y of this point are noted. Then we define $\sin A$ by

$$\sin A = y/\sqrt{x^2 + y^2} = y/r.$$

In this definition the geometrical idea of the angle is paramount. The particular units used in the measure of the angle, and therefore the particular number associated with this measurement, is unimportant as long as the angle is unchanged. It is true, for example, that

$$\begin{aligned}\sin 30^\circ &= \sin (1/12 \text{ of a revolution}) \\ &= \sin (1/3 \text{ of a right angle}) = \sin (\pi/6 \text{ radians}).\end{aligned}$$

The expression $\sin 30$ is meaningless by this definition, since 30 is merely a number and not an angle. Hence, the statement $y = \sin x$ is inadequate to define a functional relation between two numerical variables x and y , unless there is attached to it some interpretation going beyond the elementary definition of the sine. All that is needed, however, is a convention by which we agree that x is to be identified with an angle, e.g., x radians. This we shall do, and make the definition as follows.

$$\sin x = \text{sine of the number } x = \sin (x \text{ radians}).$$

The choice of the radian as the angular unit is arbitrary. Reasons why this is a convenient choice will appear as we proceed. The essential

characteristics of the function $\sin x$ would be the same if x were to refer to degrees, although the numerical value of $\sin (x \text{ degrees})$ is not the same as that of $\sin (x \text{ radians})$. Suppose, for example, two definitions were proposed.

$$\text{I. } \sin x = \sin (x \text{ rad}),$$

$$\text{II. } \sin u = \sin (u^\circ).$$

(Different letters are used so that we may know to which definition we refer.) In order to compare the functions $\sin x$ and $\sin u$ one might evaluate them for many values of x and u . If, for each value of x that is

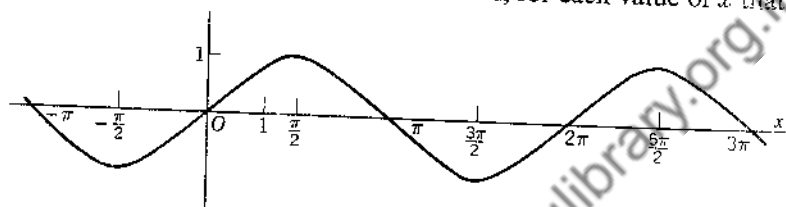


FIG. 3

used, one were to take u to be $180/\pi$ times as great, since $1 \text{ radian} = 180/\pi^\circ$, the values of $\sin x$ and $\sin u$ so obtained would agree. That is,

$$\sin (t \text{ rad}) = \sin (180/\pi \cdot t^\circ).$$

If the graph of $y = \sin x$ were drawn on xy -axes and that of $y = \sin u$ on uy -axes, and if the length representing one unit on the x -axis were $180/\pi$ times as long as that representing one unit on the u -axis, the two curves would be exactly alike.

With the identification of the numerical variable x with a number of radians a numerical function is completely defined by the expression $\sin x$. That is, for each value of the variable x , there is a definite number assigned to $\sin x$.

It is assumed that the graph of $y = \sin x$ is familiar from trigonometry. This graph is shown in Fig. 3. One unit on the x -axis, in accordance with the definition, corresponds to one radian. Some of the characteristics of this function, that have no doubt been pointed out in earlier courses and are easily seen from the graph, are worthy of emphasis.

1. The function is *periodic*, with *period* 2π . This follows from the fact that $\sin (x + 2\pi) = \sin x$. In general, a function $f(x)$ is called *periodic*, with *period* h , if $f(x + h) = f(x)$, for all values of x . If the graph of $y = \sin x$ is drawn for any interval of 2π units on the x -axis, we can then obtain the graph for an adjacent interval of 2π units by simply projecting the points already plotted 2π units to the right or left (Fig. 4).

2. The function oscillates between the extreme values 1 and -1 . The amount of oscillation from the median value is called the *amplitude*, which is 1 for the function $\sin x$.

3. The function is equal to 0 at $x = 0, \pm\pi, \pm 2\pi, \dots, \pm n\pi, \dots$.

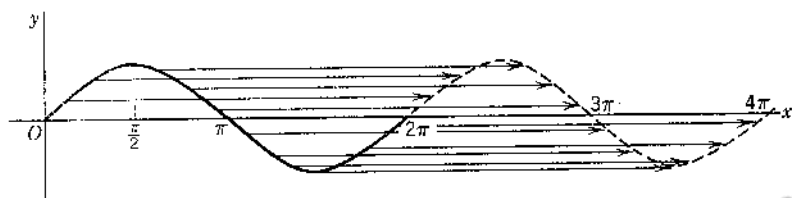


FIG. 4

4. Maxima and minima occur as follows.

Max: $x = \pi/2, 5\pi/2, 9\pi/2, \dots, -3\pi/2, -7\pi/2, \dots$,

or, in general, at $x = 2n\pi + \pi/2$, where $n = \text{any integer}$.

Min: $x = 3\pi/2, 7\pi/2, 11\pi/2, \dots, -\pi/2, -5\pi/2, \dots$,

or, in general, at $x = 2n\pi - \pi/2$, where $n = \text{any integer}$.

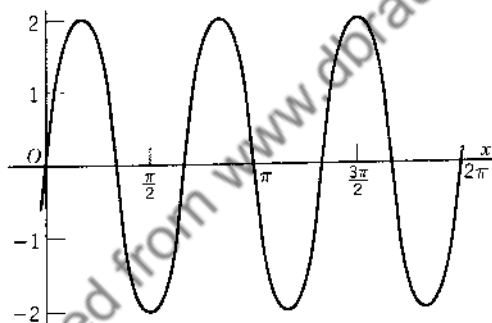


FIG. 5

The function $y = b \sin ax$ has properties similar to those of $y = \sin x$. Its amplitude is b , its period is $2\pi/a$, and there are corresponding differences in its zeros and maxima and minima. The graph of $y = 2 \sin 3x$ is shown in Fig. 5.

EXERCISES

1. Define a numerical function $y = \cos x$ in the same way that $y = \sin x$ was treated in this article. Do the same for $y = \tan x$, $y = \cot x$, $y = \sec x$, and $y = \csc x$.

2. Define $\cos x$, $\tan x$, $\cot x$, $\sec x$, and $\csc x$ in terms of $\sin x$ as defined above, by means of the relations connecting $\sin x$ with each of the other functions, which are derived by trigonometry.

3. Sketch the graph of $y = \cos x$. State the period and amplitude of this function. For what values of x is $\cos x = 0$? Where are the maximum and minimum points of the graph?

4. Sketch the graph of $y = \tan x$. What is the period of this function? The term "amplitude" is not applied to the tangent. Can you give a reason for this? For what values of x is $\tan x = 0$? What can be said about maximum and minimum points?
5. Answer Exercise 4 replacing $\tan x$ by $\cot x$.
6. Sketch the graph of $y = \csc x$, making use of the fact that $\csc x = 1/\sin x$. Discuss the question of periodicity, amplitude, zeros, maxima and minima.
7. Answer Exercise 6, replacing $\csc x$ by $\sec x$ and using $\sec x = 1/\cos x$.
8. Sketch the graph of $y = \sin 2x$.
9. Sketch the graph of $y = \cos 2x$.
10. State the amplitude and period of each of the following functions. (a) $3 \sin 4x$; (b) $2 \cos 3x$; (c) $.1 \sin 500x$; (d) $1.4 \cos x/100$; (e) $.5 \tan 2\pi x$.

4. The Continuity of $\sin x$. Limits of various algebraic functions of x were found by application of the general theorems on limits, Chapter 2,

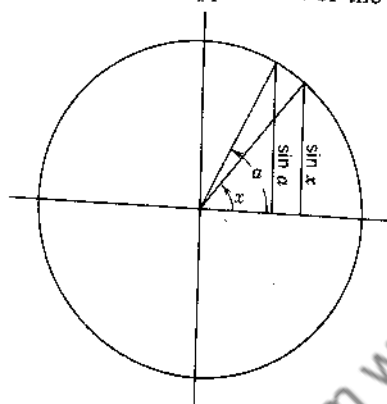


FIG. 6

Art. 9. These theorems suffice to give us the limits of functions obtained by additions, subtractions, and multiplications applied to the variable x and constants, and of many fractional expressions. They will not suffice to give us $\lim_{x \rightarrow a} \sin x$, since $\sin x$ is not

obtained from x by the algebraic operations. Nor will they tell us anything about the continuity of $\sin x$.

Since our definition of $\sin x$ rests fundamentally on geometric concepts, it is to this geometric basis that we must turn in order to examine the

fundamental nature of the function. In Fig. 6 $\sin x$ is represented by the ordinate of the point on the unit circle where the terminal side of the angle of x radians meets the circumference. Since for every value of x an ordinate is determined, we conclude that $\sin x$ is defined for all real values of x . It is also geometrically apparent that, as x tends to any particular value a , the ordinate representing $\sin x$ tends to the ordinate representing $\sin a$. That is, $\lim_{x \rightarrow a} \sin x = \sin a$. We thus conclude that $\sin x$ is continuous for all values of x .

EXERCISE

The continuity of $\sin x$ can be proved by showing that $\sin x$ can be made to differ by as little as we wish from $\sin x_0$ by taking x sufficiently close to x_0 . That is,

$$|\sin x - \sin x_0| < \epsilon \quad \text{if } |x - x_0| < \delta.$$

Show this by writing $\sin x - \sin x_0 = 2 \cos \frac{1}{2}(x + x_0) \sin \frac{1}{2}(x - x_0)$, and using the fact that $\sin \frac{1}{2}(x - x_0) < \frac{1}{2}(x - x_0)$.

5. Other Trigonometric Functions. The functions $\cos x$, $\tan x$, $\cot x$, $\csc x$, and $\sec x$, can be defined in the same way that $\sin x$ was defined in Art. 3, and the characteristics of these functions can be found from their graphs. The student was asked to supply this discussion in the exercises of Art. 3.

The question of the continuity of these functions can be examined as was that of $\sin x$ in Art. 4. In some ways it is preferable, however, to make use of the relations connecting these functions with $\sin x$ and apply the general theorems of Art. 11, Chapter 2. Thus,

(1) $\cos x = \sin(x + \pi/2)$. Hence, $\cos x$ is continuous for all values of x since $\sin(x + \pi/2)$ is continuous for all values of x . This is a simple application of the theorem at the close of Art. 12, Chapter 2.

(2) $\tan x = \sin x / \cos x$. Hence, $\tan x$ is continuous for all values except those for which $\cos x = 0$, i.e., $x = \pi/2, 3\pi/2$, etc. At these points the numerical value of $\tan x$ increases beyond all bounds. Thus, for example, as $x \rightarrow \pi/2$ $\tan x \rightarrow +\infty$ if x approaches $\pi/2$ through values less than $\pi/2$, and $\tan x \rightarrow -\infty$ if x approaches $\pi/2$ through values greater than $\pi/2$.

(3) $\cot x = \cos x / \sin x$. Hence, $\cot x$ is continuous for all values of x except those for which $\sin x = 0$, i.e., $x = 0, \pi$, etc. At these points the numerical value of $\cot x$ increases beyond all bounds.

(4) $\csc x = 1/\sin x$. Hence, $\csc x$ is continuous for all values of x except those for which $\sin x = 0$, i.e., $x = 0, x = \pi$, etc. At these points the numerical value of $\csc x$ increases beyond all bounds.

(5) $\sec x = 1/\cos x$. Hence $\sec x$ is continuous for all values of x except those for which $\cos x = 0$, i.e., $x = \pi/2, x = 3\pi/2$, etc. At these points the numerical value of $\sec x$ increases beyond all bounds.

Limits of many trigonometric expressions that are algebraic combinations of the trigonometric functions discussed above, can be found by direct application of the theorems of Chapter 2 and the results obtained in this chapter. Thus, for example, the function $\sin^2 x - 3 \tan x$ is continuous for all values of x except $x = \pi/2, x = 3\pi/2$, etc., where $\tan x$ is discontinuous. It follows that

$$\lim_{x \rightarrow \pi/4} (\sin^2 x - 3 \tan x) = \sin^2 \pi/4 - 3 \tan \pi/4 = \left(\frac{\sqrt{2}}{2}\right)^2 - 3 \cdot 1 = -2\frac{1}{2}.$$

Similarly, the function $\frac{\cos x}{\sin x - \cos x}$ is continuous for all values of x except those making the denominator, $\sin x - \cos x$, equal to 0, namely, $\pi/4, 5\pi/4, 9\pi/4$, etc. It follows that

$$\lim_{x \rightarrow \pi/2} \frac{\cos x}{\sin x - \cos x} = \frac{\cos \pi/2}{\sin \pi/2 - \cos \pi/2} = 0.$$

6. Limits of Some Discontinuous Trigonometric Functions. We have seen that many functions that are not continuous nevertheless have limits at the points of discontinuity. In fact, the derivative of a function is always such a limit, namely the limit as $h \rightarrow 0$ of $\frac{f(x_0 + h) - f(x_0)}{h}$. This function of h is discontinuous for $h = 0$, and, hence, to find its limit as $h \rightarrow 0$ we cannot insert $h = 0$. Instead we often replace $\frac{f(x_0 + h) - f(x_0)}{h}$

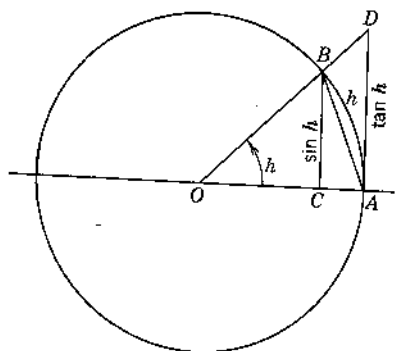


FIG. 7

by a different function, obtained by canceling out h from the numerator and denominator, which has the same limit as this one when $h \rightarrow 0$, but which is continuous at that point. We can then insert $h = 0$ in this new function to find the required limit, i.e., the derivative $f'(x_0)$. When we try to find the derivative of a trigonometric function we shall find that we cannot cancel h from numerator and denominator to obtain a new continuous function. We shall require other means of evaluating the limit of the discontinuous function $\frac{h}{\sin h}$ as $h \rightarrow 0$.

Figure 7 shows a central angle of h radians in a unit circle. From the relation, $s = \theta r$ (Art. 2), it follows that the number of units in the length of arc AB is h . Also, the number of units in the length of CB is $\sin h$, and the number of units in the length of AD is $\tan h$. Now, the areas of the triangles OAB and OAD and the sector AOB satisfy the inequality

$$\text{triangle } OAB < \text{sector } OAB < \text{triangle } OAD. \quad (6.1)$$

and

$$\text{area of triangle } OAB = \frac{1}{2} \cdot 1 \cdot CB = \frac{1}{2} \sin h,$$

$$\text{area of triangle } OAD = \frac{1}{2} \cdot 1 \cdot AD = \frac{1}{2} \tan h.$$

To find the area of the sector OAB we note the following:

$$\frac{\text{area of sector } OAB}{\text{area of the circle}} = \frac{\text{length of arc } AB}{\text{length of the circumference}};$$

i.e.,

$$\frac{\text{area of sector } OAB}{\pi} = \frac{h}{2\pi};$$

whence,

$$\text{area of sector } OAB = \frac{1}{2}h.$$

Inequality (6.1) now becomes

$$\frac{1}{2} \sin h < \frac{1}{2}h < \frac{1}{2} \tan h,$$

or, multiplying by 2,

$$\sin h < h < \tan h. \quad (6.2)$$

If we now divide this inequality by $\sin h$, which is positive if h is a small positive quantity, we get

$$1 < \frac{h}{\sin h} < \sec h.$$

As $h \rightarrow 0$, the first member remains 1, and the last member tends to 1. The second member, $h/\sin h$, must therefore have the same limit. That is,

$$\lim_{h \rightarrow 0} \frac{h}{\sin h} = 1. \quad (6.3)$$

It follows, using Theorem 4, Art. 9, Chapter 2, that

$$\lim_{h \rightarrow 0} \frac{\sin h}{h} = 1. \quad (6.4)$$

Notice that this result depends upon the fact that h is the number of radians in the central angle. If the number of degrees in the central angle is represented by u , the number of radians in the angle is $\frac{\pi u}{180}$, and this must also be the number of units in the arc AB . Hence, in place of inequality (6.2), we have

$$\sin u < \frac{\pi u}{180} < \tan u.$$

Dividing by $\sin u$, as before,

$$1 < \frac{\pi}{180} \frac{u}{\sin u} < \sec u.$$

Hence

$$\lim_{u \rightarrow 0} \frac{\pi}{180} \frac{u}{\sin u} = 1$$

and

$$\lim_{u \rightarrow 0} \frac{u}{\sin u} = \frac{180}{\pi}, \quad \lim_{u \rightarrow 0} \frac{\sin u}{u} = \frac{\pi}{180}. \quad (6.5)$$

The simplicity of the results in (6.3) and (6.4) in comparison with those in (6.5), together with the many advantages to which this simplicity leads

in later work, is the main reason for choosing the radian as the angular unit in defining the function $\sin x$.

Since the name of the variable appearing in (6.3) and (6.4) is of no consequence, we can write

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{v \rightarrow 0} \frac{\sin v}{v} = 1, \text{ etc.}$$

In words, the limit of the quotient of the sine of a variable divided by the same variable, as that variable tends to 0, is 1. Thus,

$$\lim_{2x \rightarrow 0} \frac{\sin 2x}{2x} = 1, \quad \lim_{ax \rightarrow 0} \frac{\sin ax}{ax} = 1.$$

But,

$$\lim_{x \rightarrow 0} \frac{\sin 2x}{x} \neq 1.$$

Here, the variables $2x$ and x are not the same. This limit can be found very easily as follows,

$$\lim_{x \rightarrow 0} \frac{\sin 2x}{x} = \lim_{x \rightarrow 0} 2 \cdot \frac{\sin 2x}{2x} = 2 \cdot \lim_{2x \rightarrow 0} \frac{\sin 2x}{2x} = 2.$$

Similarly,

$$\lim_{x \rightarrow 0} \frac{\sin ax}{x} = a, \quad \text{and} \quad \lim_{x \rightarrow 0} \frac{ax}{\sin x} = a.$$

As a further generalization, we find that

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\sin ax}{\sin bx} &= \lim_{x \rightarrow 0} \frac{\sin ax}{ax} \cdot \frac{bx}{\sin bx} \cdot \frac{a}{b} \\ &= \lim_{x \rightarrow 0} \frac{\sin ax}{ax} \cdot \lim_{x \rightarrow 0} \frac{bx}{\sin bx} \cdot \lim_{x \rightarrow 0} \frac{a}{b} \\ &= 1 \cdot 1 \cdot \frac{a}{b}. \end{aligned}$$

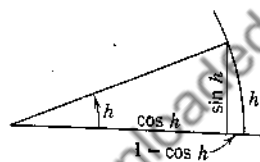


Fig. 8

In Fig. 8, we see that when h is small, $\sin h$ seems to be about the same size as h , whereas $1 - \cos h$ appears to be considerably smaller. It might be expected, therefore, that the ratio $\frac{1 - \cos h}{h}$ would tend to something less than 1 as h tends to 0. In order to find this limit, we can multiply both numerator and denominator by $1 + \cos h$. Then

$$\frac{1 - \cos h}{h} = \frac{(1 - \cos h)(1 + \cos h)}{h(1 + \cos h)} = \frac{\sin^2 h}{h(1 + \cos h)} = \frac{\sin h}{h} \cdot \frac{\sin h}{1 + \cos h}.$$

Hence,

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{1 - \cos h}{h} &= \lim_{h \rightarrow 0} \frac{\sin h}{h} \cdot \frac{\sin h}{1 + \cos h} \\ &= \lim_{h \rightarrow 0} \frac{\sin h}{h} \cdot \lim_{h \rightarrow 0} \frac{\sin h}{1 + \cos h} \\ &= 1 \cdot 0 = 0.\end{aligned}$$

When the limit of the quotient of two functions is 0, the numerator is said to *vanish to a higher order* than the denominator. Thus, $1 - \cos h$ vanishes to a higher order than h . But, by steps similar to those just taken, we can show that $1 - \cos h$ does not vanish to a higher order than h^2 . We write

$$\begin{aligned}\frac{1 - \cos h}{h^2} &= \frac{(1 - \cos h)(1 + \cos h)}{h^2(1 + \cos h)} = \frac{\sin h}{h^2(1 + \cos h)} \\ &= \frac{\sin h}{h} \cdot \frac{\sin h}{h} \cdot \frac{1}{1 + \cos h}.\end{aligned}$$

Hence,

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{1 - \cos h}{h^2} &= \lim_{h \rightarrow 0} \frac{\sin h}{h} \cdot \lim_{h \rightarrow 0} \frac{\sin h}{h} \cdot \lim_{h \rightarrow 0} \frac{1}{1 + \cos h} \\ &= 1 \cdot 1 \cdot \frac{1}{2} = \frac{1}{2}.\end{aligned}$$

EXERCISES

1. Show that $\lim_{h \rightarrow 0} \frac{\tan h}{h} = 1$.
2. Show that $1 < h/\sin h < \sec h$, if h is a small negative value.
3. Show that $\lim_{h \rightarrow 0} \frac{h^2}{\sec h - 1} = 2$.
4. Show that

$$(a) \lim_{h \rightarrow 0} \frac{3h}{\sin h} = 3.$$

$$(b) \lim_{x \rightarrow 0} \frac{1 - \cos 2x}{4x^2} = \frac{1}{2}.$$

$$(c) \lim_{h \rightarrow 0} \frac{\tan ah}{\tan bh} = \frac{a}{b}.$$

$$(d) \lim_{x \rightarrow 0} \frac{1 - \sec 2x}{x^2} = -2.$$

$$(e) \lim_{x \rightarrow 0} \frac{\sin x}{\tan x} = 1.$$

$$(f) \lim_{x \rightarrow 0} \frac{\sin^2 2x}{x \sin 3x} = \frac{4}{3}.$$

$$5. \text{ Show that } \lim_{x \rightarrow 0} \frac{(2x - x^2) \sin 2x}{x \sin x} = 4.$$

$$6. \text{ Show that } \lim_{x \rightarrow \pi} \frac{\sin x}{\pi - x} = 1. \quad [\text{Hint: Let } \pi - x = u.]$$

$$7. \text{ Show that } \lim_{x \rightarrow 0} \frac{\sin(x + \pi)}{x} = -1.$$

$$8. \text{ Show that } \lim_{x \rightarrow \pi} \frac{1 + \cos x}{\pi - x} = 0.$$

7. Differentiation of the Sine Function. With the results of Art. 6, the differentiation of $\sin x$ is now easy. Proceeding from the definition of the derivative, we have

$$D(\sin x)_{[x=x_0]} = \lim_{h \rightarrow 0} \frac{\sin(x_0 + h) - \sin x_0}{h}.$$

Since $\sin(x_0 + h) = \sin x_0 \cos h + \cos x_0 \sin h$, we have

$$\begin{aligned} \frac{\sin(x_0 + h) - \sin x_0}{h} &= \frac{\sin x_0 \cos h + \cos x_0 \sin h - \sin x_0}{h} \\ &= \frac{\cos x_0 \sin h - \sin x_0 (1 - \cos h)}{h} \\ &= \cos x_0 \frac{\sin h}{h} - \sin x_0 \frac{1 - \cos h}{h}. \end{aligned}$$

Hence,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\sin(x_0 + h) - \sin x_0}{h} &= \lim_{h \rightarrow 0} \cos x_0 \frac{\sin h}{h} - \lim_{h \rightarrow 0} \sin x_0 \frac{1 - \cos h}{h} \\ &= \cos x_0 \cdot 1 - \sin x_0 \cdot 0 \\ &= \cos x_0. \end{aligned}$$

Generalizing this result, the derived function of $\sin x$ is $\cos x$, or

$$D \sin x = \cos x. \quad (D-2)$$

The derived function of $\sin x$ would contain an additional factor if we had agreed that x should represent the number of degrees in the angle used in the definition, rather than the number of radians in this angle.

Formula D-2 depends fundamentally on the fact that $\lim_{h \rightarrow 0} \frac{\sin h}{h} = 1$, which is true if h represents the number of radians in the angle, and not true otherwise. If x is a number of degrees, so also are $x + h$ and h .

Hence, $\lim_{h \rightarrow 0} \frac{\sin h}{h} = \frac{\pi}{180}$, and the formula for the derivative of $\sin x$ is

$$D \sin x = \frac{\pi}{180} \cos x.$$

An immediate generalization of formula D-2 is obtained by use of the chain rule,

$$D_x y = D_u y \cdot D_x u, \quad \text{or} \quad D_x f(u) = D_u f(u) \cdot D_x u.$$

Applying this rule,

$$D_x \sin u = D_u \sin u \cdot D_x u,$$

$$D_x \sin u = \cos u \cdot D_x u.$$

(D-2a)

EXAMPLE. Find $D \sin (2x^2 - x)$.

SOLUTION. Using (D-2a), with $u = 2x^2 - x$, $D_x u = 4x - 1$, we have

$$D \sin (2x^2 - x) = \cos (2x^2 - x) \cdot (4x - 1).$$

8. Differentiation of Other Trigonometric Functions. The rules found in Chapter 3 for differentiating sums, products, and quotients, as well as the chain rule, are by no means restricted to use with algebraic functions. In this article we shall first apply these rules with the results of Art. 7 to obtain formulas for differentiating $\cos x$, $\tan x$, $\cot x$, $\sec x$, and $\csc x$. We shall then generalize these formulas by means of the chain rule. The illustrative examples will include other combinations of these functions.

1. Since $\cos x = \sin (x + \pi/2) = \sin u$, where $u = x + \pi/2$,

$$\begin{aligned} D \cos x &= D_u \sin u \cdot D_x u = \cos u \cdot 1 \\ &= \cos (x + \pi/2) = -\sin x. \end{aligned}$$

Thus,

$$D \cos x = -\sin x. \quad (D-3)$$

2. $\tan x = \sin x / \cos x$. Hence, applying the quotient formula, together with (D-2) and (D-3),

$$\begin{aligned} D \tan x &= \frac{\cos x \cos x - \sin x (-\sin x)}{\cos^2 x} \\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x. \end{aligned}$$

Thus,

$$D \tan x = \sec^2 x. \quad (D-4)$$

3. $\cot x = \cos x / \sin x$. Differentiating as in the preceding derivation, we obtain

$$D \cot x = -\csc^2 x. \quad (D-5)$$

4. $\sec x = 1/(\cos x)$. Hence, applying the formula

$$D_x (1/u) = -(1/u^2) D_x u,$$

we have

$$\begin{aligned} D \sec x &= -\frac{1}{\cos^2 x} D \cos x \\ &= -\frac{1}{\cos^2 x} (-\sin x), \text{ by (D-3)} \\ &= \frac{1}{\cos x} \cdot \frac{\sin x}{\cos x} = \sec x \tan x. \end{aligned}$$

Thus,

$$D \sec x = \sec x \tan x. \quad (D-6)$$

5. $\csc x = 1/\sin x$. Hence, as above,

$$\begin{aligned} D \csc x &= -\frac{1}{\sin^2 x} D \sin x \\ &= -\frac{1}{\sin^2 x} \cos x, \text{ by (D-2)} \\ &= -\frac{1}{\sin x} \cdot \frac{\cos x}{\sin x} = -\csc x \cot x. \end{aligned}$$

Thus,

$$D \csc x = -\csc x \cot x. \quad (D-7)$$

Each of the formulas (D-3) to (D-7) can now be generalized by means of the chain rule, obtaining

$$D_x \cos u = -\sin u D_x u. \quad (D-3a)$$

$$D_x \tan u = \sec^2 u D_x u. \quad (D-4a)$$

$$D_x \cot u = -\csc^2 u D_x u. \quad (D-5a)$$

$$D_x \sec u = \sec u \tan u D_x u. \quad (D-6a)$$

$$D_x \csc u = -\csc u \cot u D_x u. \quad (D-7a)$$

EXAMPLE 1. Find Dy if $y = \sin x + \cos^2 x \sin x$.

SOLUTION. By the sum formula,

$$Dy = D \sin x + D(\cos^2 x \sin x).$$

Using the product formula,

$$Dy = D \sin x + \cos^2 x D \sin x + \sin x D(\cos^2 x).$$

Now, applying (D-2), (D-3), and the power formula D-1a,

$$\begin{aligned} Dy &= \cos x + \cos^2 x \cdot \cos x + \sin x \cdot 2 \cos x (-\sin x) \\ &= \cos x + \cos^3 x - 2 \sin^2 x \cos x. \end{aligned}$$

EXAMPLE 2. Find $D\sqrt{1 + \sin x}$.

SOLUTION.

$$\begin{aligned} D\sqrt{1 + \sin x} &= D(1 + \sin x)^{\frac{1}{2}} \\ &= \frac{1}{2}(1 + \sin x)^{-\frac{1}{2}} D(1 + \sin x), \text{ by the power formula.} \\ &= \frac{1}{2\sqrt{1 + \sin x}} (0 + \cos x) \\ &= \frac{\cos x}{2\sqrt{1 + \sin x}}. \end{aligned}$$

EXAMPLE 3. Find $D(x^2 \tan^3 x)$.

SOLUTION. $D(x^2 \tan^3 x) = x^2 D(\tan^3 x) + \tan^3 x D x^2$, by use of the product formula. Differentiating $\tan^3 x$ and x^2 by the power formula,

$$\begin{aligned} D(x^2 \tan^3 x) &= x^2 \cdot 3 \tan^2 x \cdot D(\tan x) + \tan^3 x \cdot 2x \\ &= 3x^2 \tan^2 x \sec^2 x + 2x \tan^3 x, \text{ by (D-4)} \\ &= x \tan^2 x (3x \sec^2 x + 2 \tan x). \end{aligned}$$

EXAMPLE 4. Given $\sin y = y \sin x = 1$, find $D_x y$.

SOLUTION. Differentiating implicitly,

$$D_x \sin y = D_x(y \sin x) = D_x(1).$$

By the use of the product formula and formulas (D-2) and (D-2a), with the chain rule,

$$\cos y \cdot D_x y = y \cos x - \sin x \cdot D_x y = 0.$$

$$D_x y \cdot (\cos y + \sin x) = y \cos x,$$

$$D_x y = \frac{y \cos x}{\cos y + \sin x}.$$

EXAMPLE 5. Find $D_x y$, if $x = a \sin t$, and $y = b \cos t$.

SOLUTION. By the rule for differentiating parametric equations,

$$\begin{aligned} D_x y &= \frac{D_t(b \cos t)}{D_t(a \sin t)} \\ &= \frac{-b \sin t}{a \cos t}, \text{ by (D-2) and (D-3)} \\ &= -\frac{b}{a} \tan t. \end{aligned}$$

EXERCISES

Verify the results in Exercises 1-20.

- $D \sin^3 x = 3 \sin^2 x \cos x.$
- $D \sin x \cos x = \cos^2 x - \sin^2 x.$
- $D(\sin x - \sin 2x) = \cos x - 2 \cos 2x.$
- $D \tan^2 x = 2 \tan x \sec^2 x.$
- $D \sec^2 x = 2 \sec^2 x \tan x.$
- $D(\sec^2 x - \tan^2 x) = 0.$
- $D \tan x \cot x = 0.$
- $D \sec x \csc x = \sec x \csc x \cdot (\tan x - \cot x).$
- $D \cos^2 2x = -4 \sin 2x \cos 2x.$
- $D \sqrt{1 - \cos x} = \frac{1}{2} \sqrt{1 + \cos x}.$
- $D \frac{1 - \cos x}{1 + \cos x} = \frac{2 \sin x}{(1 + \cos x)^2}.$
- $D \tan(x + \pi/4) = \sec^2(x + \pi/4).$

13. $D \sin(3x - \pi/3) = 3 \cos(3x - \pi/3)$. 14. $D(x^2 \sin 2x) = 2x(\sin 2x + x \cos 2x)$.
15. $D_s \frac{\sin s}{s} = \frac{s \cos s - \sin s}{s^2}$. 16. $D_t \sqrt{\cos 2t} = -\frac{\sin 2t}{\sqrt{\cos 2t}}$.
17. $D_\theta \sin^2(\pi - \theta) = 2 \sin \theta \cos \theta$.
18. $D_r(\sin 2r \cos r - \cos 2r \sin r) = \cos r$.
19. $D_z(z^2 \tan z/2) = 2z \tan \frac{z}{2} + \frac{z^2}{2} \sec^2 z/2$.
20. $D_\theta \sin(\theta + \pi) \sin(\theta - \pi) = 2 \sin \theta \cos \theta$.

Find $f''(x)$ from each of the functions in Exercises 21–28.

21. $f(x) = \sin x$. 22. $f(x) = \cos x$.
23. $f(x) = \sin ax$. 24. $f(x) = \tan x$.
25. $f(x) = \sin^2 3x$. 26. $f(x) = x \cos x$.
27. $f(x) = a \sin x + b \cos x$. 28. $f(x) = a \sin(bx + c)$.

29. Find $D_x y$ if $y = \sin(x + y)$.

30. Find $D_x y$ if $x = \sin y$.

31. Find $D_x y$ if $\sin(x + y) + \cos(x - y) = \frac{1}{2}$.

32. Find $D_x y$ if $x \tan y - xy = 1$.

33. Find $D_x y$ if $x = a \cos^3 t$, $y = a \sin^3 t$.

34. Find $D_x y$ if $x = a(\theta - \sin \theta)$, $y = a(1 - \cos \theta)$.

35. Find $D_x y$ if $x = (a \cos t)/t$, $y = (a \sin t)/t$.

36. Find the slope of the graph of $y = \sin x$ at $x = 0$, and at $x = \pi$.

37. Find the slope of the graph of $y = \tan x$ at $x = 0$, and at $x = \pi/4$.

38. Find the slope of the graph of $y = \frac{1}{2} \cos 2x$ at each crossing of the x -axis.

39. Find the slope of the graph of $y = x \sin x$ at $x = 0$, and at $x = \pi/2$.

40. Find a maximum point of the graph of $y = \sin x + \cos x$.

9. The Functional Relation $y = \sin^{-1} x$. If we wish to consider y as the function of x that is expressed by the equation $\sin y = x$, we write

$$y = \sin^{-1} x, \quad \text{or} \quad y = \arcsin x.$$

Either of these symbols is read "inverse sine x ."

Since $y = \sin^{-1} x$ and $\sin y = x$ express the same relationship between x and y , it is clear that x is a function of y that has all the properties described in Art. 3, with the variables interchanged. The graph of this functional relation is shown in Fig. 9. This curve represents x as a single-valued function of y . If we wish to change our point of view so as to look upon y as a function of x , we observe that the range of values permitted x is $-1 \leq x \leq 1$. Furthermore, if y is to be a single-valued function of x , it is necessary to restrict the values of y so that only one value occurs for any given value of x . In order to do this, it is customary to use only the *principal value* of the angle $\sin^{-1} x$, defined as follows.

DEFINITION. For any x in the range $-1 \leq x \leq 1$, the principal value of $\sin^{-1} x$ is its numerically smallest value.

Thus, when restricted to the principal value, $\sin^{-1} 1 = \pi/2$, $\sin^{-1} (-1) = -\pi/2$, $\sin^{-1} \frac{1}{2} = \pi/6$, $\sin^{-1} (-\frac{1}{2}) = -\pi/6$, $\sin^{-1} 0 = 0$, etc.

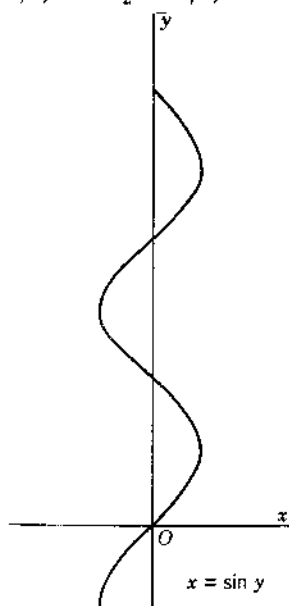


FIG. 9

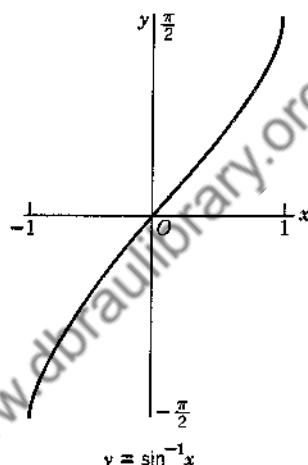


FIG. 10

Figure 10 shows the graph of $y = \sin^{-1} x$, where $\sin^{-1} x$ is restricted to its principal value. This part of the graph is called the principal branch of the curve in Fig. 9. If one wishes to refer to any other branch, that fact must be stated. Figure 11 shows the graph of the branch obtained by taking the values of $\sin^{-1} x$ that lie between $\pi/2$ and $3\pi/2$. Hereafter, $\sin^{-1} x$ will refer to the principal value unless stated otherwise.

From Fig. 10 it appears that the function $\sin^{-1} x$ always increases as x increases. It has its minimum value $-\pi/2$ when $x = -1$, and its maximum value $\pi/2$ when $x = 1$. These points are not associated with tangents having a 0 slope, but occur at the ends of the interval -1 to 1 . This function is continuous and has a derivative everywhere except at the end points of the interval of definition. At these points the quotient

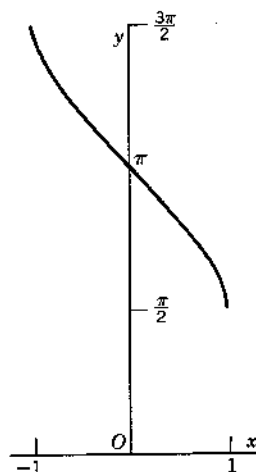


FIG. 11

$\frac{f(x_0 + h) - f(x_0)}{h}$ tends to infinity as h tends to 0 (see Exercise 10, Art. 10).

In order to differentiate $y = \sin^{-1} x$ at any other point in the range, we apply to $x = \sin y$ the formula for differentiating an inverse function, viz., $D_x y = 1/D_y x$. Thus, if $x = \sin y$,

$$D_y x = \cos y,$$

and

$$D_x y = 1/\cos y.$$

Now,

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2}.$$

Therefore,

$$D \sin^{-1} x = 1/\sqrt{1 - x^2}.^*$$
(D-8)

Generalizing, by the chain rule,

$$D_x \sin^{-1} u = 1/\sqrt{1 - u^2} \cdot D_x u.$$
(D-8a)

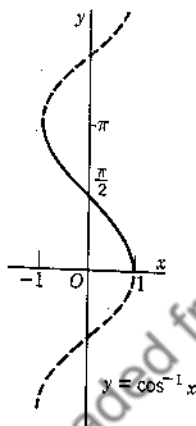


FIG. 12

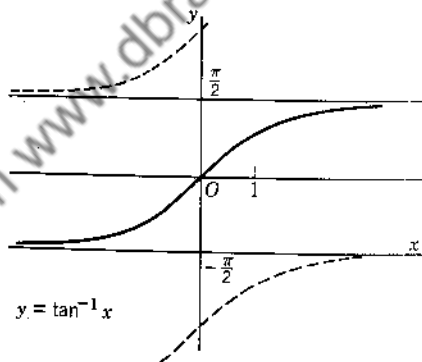


FIG. 13

10. Other Inverse Trigonometric Functions. The functions $\cos^{-1} x$, $\tan^{-1} x$, $\cot^{-1} x$, $\sec^{-1} x$, and $\csc^{-1} x$ can be treated in the same way that we dealt with $\sin^{-1} x$ in Art. 9. Their graphs are shown in Figs. 12-16, with the graph of the principal value in each case indicated by the solid line.†

* The double sign $\pm$ is not needed, since the principal value of $\sin^{-1} x$ lies between $-\pi/2$ and $\pi/2$, in which range the cosine is always positive. If we were considering the branch of the function $\sin^{-1} x$ between $\pi/2$ and $3\pi/2$, a negative sign would be needed, since in this range the cosine is always negative.

† The definitions of principal values are somewhat arbitrary and in other books may be found to differ from the ones adopted here. Any one of many choices is possible as long as it provides a single value for each value of x in the range.

The definitions of the principal values of $\tan^{-1} x$, $\cot^{-1} x$, and $\csc^{-1} x$ are similar to that of the principal value of $\sin^{-1} x$. For each, the principal value is the numerically smallest value. The principal value of $\cos^{-1} x$ is the smallest positive value, and the principal value of $\sec^{-1} x$ is defined in the same way.

Formulas for differentiating these functions are found as in Art. 9. These are derived below for $\cos^{-1} x$ and $\tan^{-1} x$.

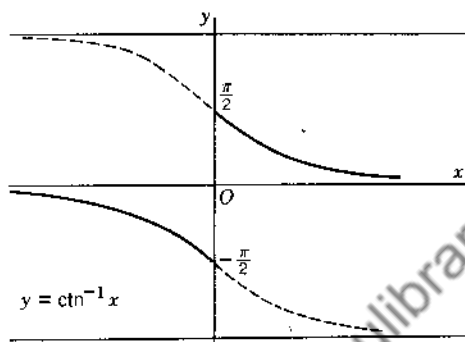


FIG. 14

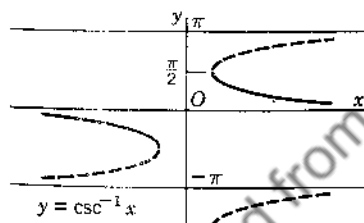


FIG. 15

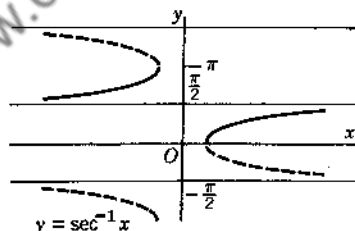


FIG. 16

1. If

$$y = \cos^{-1} x, \quad x = \cos y.$$

Then

$$D_y x = -\sin y, \quad \text{and} \quad D_x y = -\frac{1}{\sin y}.$$

Since

$$\sin y = \sqrt{1 - \cos^2 y} = \sqrt{1 - x^2}$$

$$D_x y = D \cos^{-1} x = -\frac{1}{\sqrt{1 - x^2}}. \quad (D-9)$$

Generalizing, by the chain rule,

$$D_x \cos^{-1} u = -\frac{1}{\sqrt{1 - u^2}} D_x u. \quad (D-9a)$$

2. If

$$y = \tan^{-1} x, \quad x = \tan y.$$

Then,

$$D_y x = \sec^2 y, \quad \text{and} \quad D_x y = \frac{1}{\sec^2 y}.$$

Since

$$\sec^2 y = 1 + \tan^2 y = 1 + x^2,$$

$$D_x y = D \tan^{-1} x = \frac{1}{1 + x^2}. \quad (D-10)$$

Generalizing, by the chain rule,

$$D_x \tan^{-1} u = \frac{1}{1 + u^2} D_x u. \quad (D-10a)$$

The derivation of formulas for the derived functions of $\cot^{-1} x$, $\sec^{-1} x$, and $\csc^{-1} x$ is left to the student. The results are as follows.

$$D \cot^{-1} x = -\frac{1}{1 + x^2}. \quad (D-11)$$

$$D \sec^{-1} x = \left| \frac{1}{x\sqrt{x^2 - 1}} \right|. \quad (D-12)$$

$$D \csc^{-1} x = -\left| \frac{1}{x\sqrt{x^2 - 1}} \right|. \quad (D-13)$$

The generalizations of these formulas are:

$$D_x \cot^{-1} u = -\frac{1}{1 + u^2} D_x u. \quad (D-11a)$$

$$D_x \sec^{-1} u = \left| \frac{1}{u\sqrt{u^2 - 1}} \right| D_x u. \quad (D-12a)$$

$$D_x \csc^{-1} u = -\left| \frac{1}{u\sqrt{u^2 - 1}} \right| D_x u. \quad (D-13a)$$

EXAMPLE 1. Find $D \sin^{-1} x^2$.

SOLUTION. Using formula D-8a,

$$D \sin^{-1} x^2 = \frac{1}{\sqrt{1 - (x^2)^2}} D(x^2) = \frac{2x}{\sqrt{1 - x^4}}.$$

EXAMPLE 2. Find $D \tan^{-1}(x/a)$.

SOLUTION. Using formula D-10a,

$$\begin{aligned} D \tan^{-1} \left(\frac{x}{a} \right) &= \frac{1}{1 + (x/a)^2} D \left(\frac{x}{a} \right) \\ &= \frac{1/a}{1 + x^2/a^2} = \frac{a}{a^2 + x^2}. \end{aligned}$$

EXAMPLE 3. Find $D \cos^{-1}(x^3 - 3x)$.

SOLUTION. Using formula D-9a,

$$\begin{aligned} D \cos^{-1}(x^3 - 3x) &= - \frac{1}{\sqrt{1 - (x^3 - 3x)^2}} D(x^3 - 3x) \\ &= - \frac{1}{\sqrt{1 - x^6 + 6x^4 - 9x^2}} (3x^2 - 3) \\ &= - \frac{3(x^2 - 1)}{\sqrt{1 - 9x^2 + 6x^4 - x^6}}. \end{aligned}$$

EXERCISES

Verify the results in Exercises 1-9.

$$1. D \sin^{-1} \left(\frac{x}{a} \right) = \frac{1}{\sqrt{a^2 - x^2}}.$$

$$2. D \csc^{-1} \left(\frac{x}{a} \right) = - \frac{a}{a^2 + x^2}.$$

$$3. D \csc^{-1} \left(\frac{1}{x} \right) = \frac{1}{\sqrt{1 - x^2}}.$$

$$4. D \tan^{-1} \left(\frac{a}{x} \right) = - \frac{a}{a^2 + x^2}.$$

$$5. D \sin^{-1} \sqrt{x} = \frac{1}{2\sqrt{x}\sqrt{1-x}}.$$

$$6. D x^2 \sin^{-1} x = 2x \sin^{-1} x + \frac{x^2}{\sqrt{1-x^2}}.$$

$$7. D_t \tan^{-1} \frac{c+t}{1-ct} = \frac{1}{1+t^2}.$$

$$8. D_s(s \sin^{-1} \sqrt{1-s^2}) = \cos^{-1} s - \frac{s}{\sqrt{1-s^2}}.$$

$$9. D_t \left[t\sqrt{a^2 - t^2} + a^2 \sin^{-1} \left(\frac{t}{a} \right) \right] = 2\sqrt{a^2 - t^2}.$$

10. Show that

$$\frac{\sin^{-1}(1-h) - \sin^{-1} 1}{-h} \rightarrow \infty \quad \text{as } h \rightarrow 0.$$

$$\begin{aligned} \left[\text{Hint: } \sin^{-1}(1-h) - \sin^{-1} 1 &= \sin^{-1}(1-h) - \pi/2 \right. \\ &= -\cos^{-1}(1-h). \end{aligned}$$

$$\text{Let } \cos^{-1}(1-h) = A, \text{ and use the fact that } \lim_{A \rightarrow 0} \frac{1 - \cos A}{A} = 0.]$$

11. Applications. It is obvious, of course, that the applications of derivatives that have been mentioned in connection with algebraic functions are by no means restricted to those functions. As we learn how to differentiate more and more functions we shall be able to solve tangent problems, rate problems, and problems in maxima and minima associated with those functions. We shall now consider some examples of these familiar problems in which trigonometric functions and their inverses are involved.

EXAMPLE 1. Plot the graph of $y = 2 \cos x - \sin 2x$. Find the slope formula and the critical points.

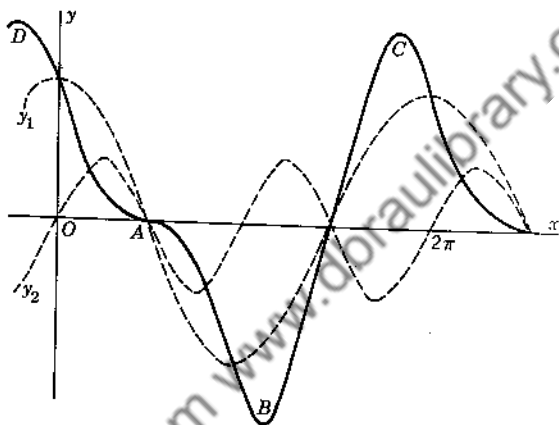


FIG. 17

SOLUTION. The graph of a trigonometric function made up of a sum or a difference, like this one, is best drawn with the help of two auxiliary curves: here, $y_1 = 2 \cos x$, and $y_2 = \sin 2x$. After drawing these on the same axes, we subtract the ordinates $y_1 - y_2$ in order to obtain the graph of $y = 2 \cos x - \sin 2x$ (Fig. 17).

Differentiating, we get

$$y' = -2(\sin x + \cos 2x).$$

Hence, the slope at $x = 0$, $y = 2$ is $-2(0 + 1) = -2$; the slope at $x = \pi/2$, $y = 0$ is $-2(1 - 1) = 0$; etc. The critical points are found by setting $y' = 0$. Thus,

$$-2(\sin x + \cos 2x) = 0.$$

Dividing by 2 and replacing $\cos 2x$ by its equal $1 - 2 \sin^2 x$, we have

$$2 \sin^2 x - \sin x - 1 = 0.$$

Factoring.

$$(2 \sin x + 1)(\sin x - 1) = 0,$$

whence

$$\sin x = 1 \quad \text{or} \quad -\frac{1}{2}.$$

The points at which $\sin x = 1$ are $x = \pi/2, y = 0; x = 5\pi/2, y = 0;$ etc. The point $(\pi/2, 0)$ is seen to be a level inflection point, and due to the periodic nature of the curve, so also is $(5\pi/2, 0)$, and so on. One point at which $\sin x = -\frac{1}{2}$ is $x = 7\pi/6, y = -\frac{3}{2}\sqrt{3}$, and this is the minimum point B . Another point at which $\sin x = -\frac{1}{2}$ is $x = 11\pi/6, y = \frac{3}{2}\sqrt{3}$, which is a maximum point C . Because of the periodicity of the curve, there are minimum points at $2\pi, 4\pi, \dots, 2n\pi$ units to the right and left of B , and maximum points at $2\pi, 4\pi, \dots, 2n\pi$ units to the right and left of C . One of the latter is point D .

It is easily verified that the second derivative is equal to 0 at the point A .

$$y'' = -2 \cos x + 4 \sin 2x.$$

At $x = \pi/2, y'' = -2 \cos \pi/2 + 4 \sin \pi = 0$.

EXAMPLE 2. An airplane 4000 ft high is flying horizontally at a constant speed on a line that passes directly over an observer. When the airplane is 3000 ft from the point directly overhead, the observer finds that he must revolve his searchlight at the rate of 3° per sec in order to keep it trained on the airplane. What is the speed of the airplane?

SOLUTION. From Fig. 18 we write

$$x = 4000 \tan \theta. \quad (11.1)$$

We think of both x and θ as functions of a time variable, t seconds, related by Eq. 11.1. Differentiating with respect to t ,

$$D_t x = 4000 \sec^2 \theta D_t \theta. \quad (11.2)$$

When $x = 3000$, $\tan \theta = \frac{3}{4}$, and $\sec \theta = \frac{5}{4}$. The angular rate of 3° per sec corresponds to $D_t \theta = 3(\pi/180) = \pi/60$. Putting this value and $\sec \theta = \frac{5}{4}$ in (11.2), we have

$$Dx = 4000 \frac{25}{16} \cdot \frac{\pi}{60} = 327\frac{1}{4}.$$

Since the distances are measured in feet and $D_t \theta$ corresponds to radians per second, the speed is in feet per second. The speed of the airplane is therefore $327\frac{1}{4}$ ft per sec, which is the same as $223\frac{1}{8}$ mph.

EXAMPLE 3. A picture 5 ft in height is hung on a wall so that its bottom edge is 9 ft above the floor. At what distance should a viewer stand so

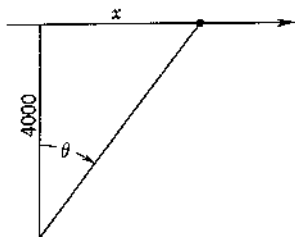


FIG. 18

that the height of the picture will subtend the greatest angle at his eye, which is 5 ft above the floor?

SOLUTION. The situation described in the problem is illustrated in Fig. 19. Our problem is to find the value of x that makes the angle θ a maximum. We note that θ is the difference of the two angles AOC and BOC , and that $\text{ctn } AOC = x/9$ and $\text{ctn } BOC = x/4$. Hence,

$$\theta = \text{ctn}^{-1} x/9 - \text{ctn}^{-1} x/4.$$

Differentiating, by use of formula D-11a,

$$D_x \theta = -\frac{1/9}{1 + x^2/9^2} + \frac{1/4}{1 + x^2/4^2}.$$

Setting this derived function equal to 0, in order to find the value of x that makes θ maximum, we get

$$\frac{4}{16 + x^2} - \frac{9}{81 + x^2} = 0.$$

Clearing fractions,

$$4 \cdot 81 + 4x^2 - 9 \cdot 16 - 9x^2 = 0,$$

whence

$$5x^2 = 180 \text{ and } x = 6.$$

Thus, the distance for the maximum value of θ is 6 ft.

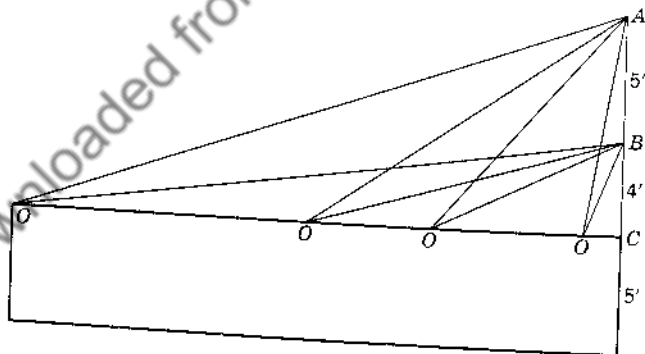


FIG. 20

The fact that this value corresponds to a maximum rather than a minimum is easily seen by reasoning as we did in Chapter 5. If x is small the angle θ is small, and when x is large the angle θ is again small, with larger values between. As x increases, θ must therefore increase to a maximum and then decrease again. This is illustrated in Fig. 20.

EXAMPLE 4. Find the dimensions of the cylinder of greatest volume that can be inscribed in a sphere of radius 8.

SOLUTION. This problem is like Exercise 19, Art. 4, Chapter 5. It is taken up here to illustrate the use of an angle as a variable in maximum

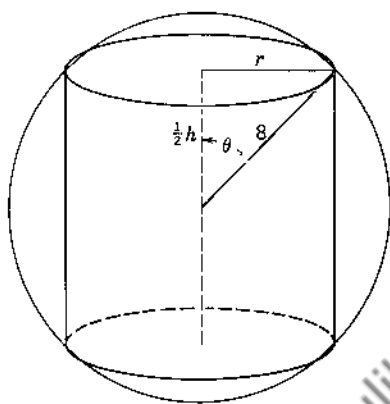


FIG. 21

and minimum problems of this kind. We have to maximize the volume $V = \pi r^2 h$. From Fig. 21,

$$r = 8 \sin \theta, \quad h = 2 \cdot 8 \cos \theta.$$

Hence,

$$\begin{aligned} V &= \pi \cdot 64 \cdot \sin^2 \theta \cdot 16 \cdot \cos \theta \\ &= 1024\pi \sin^2 \theta \cos \theta. \end{aligned}$$

Differentiating,

$$D_{\theta} V = 1024\pi (2 \sin \theta \cos^2 \theta - \sin^3 \theta).$$

Setting the derived function equal to 0, and dividing out the factor $1024\pi \sin \theta$, which cannot be equal to 0, if V is maximum,

$$2 \cos^2 \theta - \sin^2 \theta = 0.$$

Transposing and dividing by $\cos^2 \theta$,

$$\frac{\sin^2 \theta}{\cos^2 \theta} = 2 \quad \text{or} \quad \tan^2 \theta = 2,$$

whence

$$\tan \theta = \sqrt{2}, \quad \sin \theta = \sqrt{2}/\sqrt{3}, \quad \cos \theta = 1/\sqrt{3}.$$

Putting these values in the expressions for r and h , we have for a maximum volume

$$r = 8\sqrt{2}/\sqrt{3}, \quad h = 16/\sqrt{3}.$$

EXERCISES

- Sketch the graph of $y = x + 2 \sin x$ in the interval $0 \leq x \leq 2\pi$. State the coordinates of the maximum and minimum points, if any. At what angle does the curve cut the line $y = x$ at $x = \pi$?
- Sketch the graph of $y = \tan x - x$ in the interval $-\pi/2 < x < \pi/2$. Show that the point $(0, 0)$ is a level inflection point.
- Find maximum and minimum points of the graph of $y = \tan x - 2x$, in the interval $-\pi/2 < x < \pi/2$, and sketch the curve.
- Find maximum and minimum points of the graph of $y = 2x - \sin^{-1} x$, and sketch the curve.
- A revolving beacon is situated 3000 ft off a straight shore. If the beacon turns at the rate of 2 complete revolutions per min, how fast does the beam sweep along the shore at the nearest point? at a point 4000 ft from the nearest point? [Hint: $D_t \theta = 4\pi$.]
- A ladder 20 ft long leans against a wall. The lower end is being pulled away from the wall in such a way that the ladder turns at the rate of 12° per sec. How fast is the top of the ladder sliding down when the inclination of the ladder is 60° ? How fast is the lower end being moved away at this time?
- If, in Exercise 6, the lower end is moving away at 3 ft per sec, how fast is the ladder turning?
- If the angle of elevation of the sun is decreasing at $\frac{1}{4}$ rad per hr, how fast is the shadow on a horizontal surface cast by a vertical pole 60 ft long lengthening at the moment when the sun's elevation is $\pi/6$?
- A vertical pole stands 20 ft west of a tall building on which it casts a shadow. If the sun is going down at the rate of $\frac{1}{4}$ rad per hr, how fast is the shadow of the pole lengthening when the elevation of the sun is 30° ? 10° ?
- If a body moves in a straight line so that its distance s is a function of t , given by $s = a \sin ct + b \cos ct$, in which a , b , and c are constants, show that the acceleration is always proportional to s .
- A body moves in a straight line so that its distance s inches is given by the equation $s = \sin(\pi t/12)$, in which t is the time in seconds. Find its velocity at $t = 0$, $t = 2$, $t = 3$, $t = 6$, $t = 12$. Describe its motion during the first 12 sec.
- Two sides of a triangle are 12 and 15 ft long, respectively. How fast is the third side changing if the angle between the given sides is 30° and is increasing at the rate of 2° per sec? [Hint: Use the law of cosines.]
- How fast is the area changing in Exercise 13?
- A crank OB , 9 in. long, revolves about O at the rate of 360 revolutions per min. A connecting rod AB , 24 in. long, is attached to a piston rod at A . Find the speed of the point A when the crank is in the position shown in Fig. 22. [Hint: Use the law

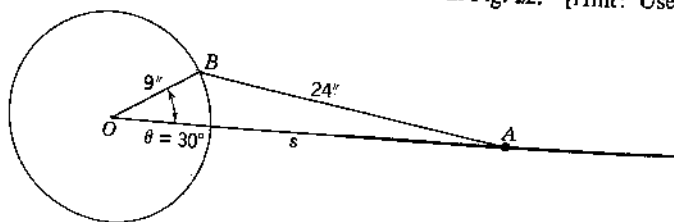


FIG. 22

of cosines, $24^2 = 9^2 + s^2 - 18s \cos \theta$, and differentiate implicitly. Now, set $\theta = 30^\circ$, solve the triangle for s , and then find $D_s s$.]

15. Solve Exercise 23, Art. 4, Chapter 5, by expressing the time as a function of the angle at A .

16. Solve Exercise 5, Art. 5, Chapter 5, by expressing the volume and area as functions of the vertex angle θ .

17. Solve Exercise 16, Art. 5, Chapter 5, by expressing the lateral area in terms of a suitable angle.

18. Solve Exercise 24, Art. 5, Chapter 5, by expressing the volume as a function of the vertex angle θ (Fig. 23).

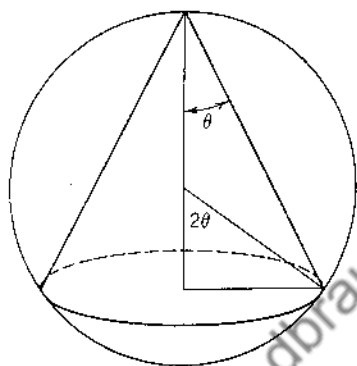


FIG. 23

12. Inverse Differentiation. Just as we reversed the formulas for differentiating algebraic functions to obtain formulas for inverse derivatives, so we can obtain inverse differentiation formulas from the derivatives found in this chapter. The resulting formulas fall in two classes: those that are inverse derivatives of trigonometric functions and result in trigonometric functions, and those that are inverse derivatives of algebraic functions and result in inverse trigonometric functions. These formulas are listed below. Each can be proved by differentiation. This will be left to the student. The numbering of the formulas corresponds to that of the differentiation formulas; e.g., (I-2) corresponds to (D-2).

$$D^{-1} \cos x = \sin x + C; \text{ or } \int \cos x \, dx = \sin x + C. \quad (I-2)$$

$$D^{-1} \sin x = -\cos x + C; \text{ or } \int \sin x \, dx = -\cos x + C. \quad (I-3)$$

$$D^{-1} \sec^2 x = \tan x + C; \text{ or } \int \sec^2 x \, dx = \tan x + C. \quad (I-4)$$

$$D^{-1} \csc^2 x = -\cot x + C; \text{ or } \int \csc^2 x \, dx = -\cot x + C. \quad (I-5)$$

$$D^{-1} \sec x \tan x = \sec x + C; \text{ or } \int \sec x \tan x \, dx = \sec x + C. \quad (I-6)$$

$$D^{-1} \csc x \cot x = -\csc x + C; \text{ or } \int \csc x \cot x \, dx = -\csc x + C. \quad (I-7)$$

If $|x| < 1$,

$$D^{-1}(1/\sqrt{1-x^2}) = \sin^{-1} x + C;$$

or

$$\int (1/\sqrt{1-x^2}) dx = \sin^{-1} x + C. \quad (I-8)$$

If $|x| < 1$,

$$D^{-1}(1/\sqrt{1-x^2}) = -\cos^{-1} x + C;$$

or

$$\int (1/\sqrt{1-x^2}) dx = -\cos^{-1} x + C. \quad (I-9)$$

$$D^{-1}[1/(1+x^2)] = \tan^{-1} x + C;$$

or

$$\int [1/(1+x^2)] dx = \tan^{-1} x + C. \quad (I-10)$$

$$D^{-1}[1/(1+x^2)] = -\operatorname{ctn}^{-1} x + C;$$

or

$$\int [1/(1+x^2)] dx = -\operatorname{ctn}^{-1} x + C. \quad (I-11)$$

If $|x| > 1$,

$$\begin{aligned} D^{-1}[1/(x\sqrt{x^2-1})] &= \sec^{-1} x + C, \text{ if } x > 0, \\ &= -\sec^{-1} x + C, \text{ if } x < 0; \end{aligned}$$

or

$$\begin{aligned} \int [1/(x\sqrt{x^2-1})] dx &= \sec x + C, \text{ if } x > 1, \\ &= -\sec^{-1} x + C, \text{ if } x < 1. \end{aligned} \quad (I-12)$$

If $|x| > 1$,

$$\begin{aligned} D^{-1}[1/(x\sqrt{x^2-1})] &= -\csc^{-1} x + C, \text{ if } x > 0, \\ &= \csc^{-1} x + C, \text{ if } x < 0; \end{aligned}$$

or

$$\begin{aligned} \int [1/(x\sqrt{x^2-1})] dx &= -\csc^{-1} x + C, \text{ if } x > 0, \\ &= \csc^{-1} x + C, \text{ if } x < 0. \end{aligned} \quad (I-13)$$

Formulas I-9, I-11, and I-13 are alternatives for I-8, I-10, and I-12, respectively, and are seldom used. Three further formulas will be added which are generalizations of formulas I-8, I-10, and I-12, respectively. They can be easily verified (see Exercises 13 and 14 below).

If $x^2 < a^2$,

$$D^{-1}(1/\sqrt{a^2-x^2}) = \sin^{-1} x/a + C;$$

or

$$\int (1/\sqrt{a^2-x^2}) dx = \sin^{-1} x/a + C. \quad (I-8')$$

or

$$D^{-1}[1/(a^2+x^2)] = 1/a \tan^{-1} x/a + C;$$

$$\int [1/(a^2+x^2)] dx = 1/a \tan^{-1} x/a + C. \quad (I-10')$$

If $x^2 > a^2$,

$$\begin{aligned} D^{-1}[1/(x\sqrt{x^2 - a^2})] &= 1/a \sec^{-1} x/a + C, \text{ if } x > 0, \\ &= -1/a \sec^{-1} x/a + C, \text{ if } x < 0; \end{aligned}$$

or

$$\begin{aligned} \int [1/(x\sqrt{x^2 - a^2})] &= 1/a \sec^{-1} x/a + C, \text{ if } x > 0, \\ &= -1/a \sec^{-1} x/a + C, \text{ if } x < 0. \quad (I-12') \end{aligned}$$

Corresponding to each of the above inverse derivative formulas is a more general formula obtained from the more general differentiation formula by the chain rule. Thus, from (D-2a) we see that

$$D^{-1}[\cos u D_x u] = \sin u + C; \text{ or } \int \cos u (D_x u) dx = \sin u + C.$$

We shall not list these general formulas here. Each of the formulas of the above list can be extended so that x can be replaced by a function of x , provided that the derivative of this function is present as a factor. This will be illustrated in the examples below. The general power formula I-1a will also be used frequently. For example,

$$\begin{aligned} D^{-1}(\sin^2 x \cos x) &= D^{-1}[u^2 D_x u], \text{ where } u = \sin x \\ &= u^3/3 + C = \sin^3 x/3 + C. \end{aligned}$$

Formulas I-8 to I-13 extend our knowledge of the primitive functions of algebraic functions to include several types that could not be handled in Chapter 6. Hence, our ability to find definite integrals is also extended to several new kinds of algebraic functions. What is most remarkable about these formulas, however, is the group of relationships between algebraic functions and transcendental functions that they exhibit. Transcendental functions are, by definition, functions that cannot be obtained from algebraic expressions by means of the operations of addition, subtraction, multiplication, division, involution, and evolution—the so-called algebraic operations. It is found, however, that certain algebraic functions have primitive functions that are transcendental. For example, $\tan^{-1} x$ is a primitive function of $1/(1 + x^2)$. In Chapter 11 we shall see that other transcendental functions, logarithmic and exponential functions, have a similar connection with algebraic functions.

The following examples illustrate the use of the formulas of this article. Each of these examples and the exercises that follow could be stated equally correctly in three ways. For example, the first one might be stated: evaluate $D^{-1} \sin 2x$, evaluate $\int \sin 2x dx$, or solve the differential equation $D_x y = \sin 2x$.

EXAMPLE 1. Evaluate $D^{-1} \sin 2x$.

SOLUTION. Formula I-3 could be applied if the derivative of $2x$ were present as a factor. This derivative is 2 and it can be introduced if

balanced by a division by 2. We have already observed, in Chapter 6, that a constant factor can be moved from one side of the inverse-derivative symbol to the other. Hence, we proceed as follows.

$$D^{-1} \sin 2x = D^{-1} \sin 2x \cdot 2 \cdot \frac{1}{2} = \frac{1}{2} D^{-1} \sin 2x \cdot 2.$$

This has the form $\frac{1}{2} D^{-1} [\sin u \cdot D_x u]$, with $u = 2x$, and hence the result is $-\frac{1}{2} \cos 2x + C$.

EXAMPLE 2. Evaluate $D^{-1} \sec^2 (x^2 - x) \cdot (2x - 1)$.

SOLUTION. The factor, $2x - 1$, is the derivative of $x^2 - x$. Hence,

$$\begin{aligned} D^{-1} \sec^2 (x^2 - x) \cdot (2x - 1) &= D^{-1} [\sec^2 u \cdot D_x u], \quad u = x^2 - x \\ &= \tan u + C \\ &= \tan (x^2 - x) + C. \end{aligned}$$

EXAMPLE 3. Evaluate $D^{-1} \cos^3 x \sin x$.

SOLUTION. Since $D \cos x = -\sin x$,

$$\begin{aligned} D^{-1} \cos^3 x \sin x &= -D^{-1} \cos^3 x \cdot D \cos x \\ &= -D^{-1} [u^3 \cdot D_x u], \quad \text{where } u = \cos x \\ &= -\frac{u^4}{4} + C = -\frac{\cos^4 x}{4} + C. \end{aligned}$$

EXAMPLE 4. Evaluate $D^{-1} \tan^2 x \sec^2 x$.

SOLUTION. Since $D \tan x = \sec^2 x$, $D^{-1} \tan^2 x \sec^2 x = D^{-1} [u^2 \cdot D_x u]$, where $u = \tan x$

$$= \frac{u^3}{3} + C = \frac{\tan^3 x}{3} + C.$$

EXAMPLE 5. Evaluate $D^{-1} \frac{2}{4 + x^2}$.

SOLUTION. By (I-10'), $D^{-1} \frac{1}{4 + x^2} = \frac{1}{2} \tan^{-1} \frac{x}{2}$. Hence,

$$D^{-1} \frac{2}{4 + x^2} = 2 D^{-1} \frac{1}{4 + x^2} = \tan^{-1} \frac{x}{2} + C.$$

EXERCISES

Find the inverse derivatives in Exercises 1-20.

1. $D^{-1} \cos 3x = \frac{1}{3} \sin 3x + C$.
2. $D^{-1} \cos^3 x \sin x = -\frac{1}{4} \cos^4 x + C$.
3. $D^{-1} \sec^3 x \tan x = \frac{1}{3} \sec^3 x + C$. [Let $u = \sec x$.]

4. $D^{-1} \csc^3 x \csc^2 x = -\frac{1}{3} \csc^4 x + C.$

5. $D^{-1} \sin^2 2x \cos 2x = \frac{1}{8} \sin^3 2x + C.$

6. $D^{-1} \sec^4 x = \tan x + \frac{1}{3} \tan^3 x + C.$ [Write $\sec^4 x = (1 + \tan^2 x) \sec^2 x.$]

7. $D^{-1} \tan^4 x = \frac{1}{3} \tan^3 x - \tan x + x + C.$

[Write $\tan^4 x = \tan^2 x \sec^2 x - \sec^2 x + 1.$]

8. $D^{-1} \sin^3 x = -\cos x + \frac{1}{3} \cos^3 x + C.$ [Write $\sin^3 x = (1 - \cos^2 x) \sin x.$]

9. $D^{-1} \cos^3 x = \sin x - \frac{1}{3} \sin^3 x + C.$

10. $D^{-1} \sqrt{1 - \sin x} \cos x = -\frac{2}{3} (1 - \sin x)^{3/2} + C.$

11. $\int 2x/\sqrt{1-x^2} dx = \sin^{-1} x^2 + C.$

12. $\int 1/\sqrt{4-x^2} dx = \sin^{-1} (x/2) + C.$

13. $\int a/(a^2-x^2) dx = \tan^{-1} (x/a) + C.$

14. $\int 1/\sqrt{a^2-x^2} dx = \sin^{-1} (x/a) + C.$

15. $\int 3/(2x\sqrt{4x^2-1}) dx = \frac{3}{2} \sec^{-1} 2x + C.$

16. $\int (\tan^{-1} x)/(1+x^2) dx = \frac{1}{2} (\tan^{-1} x)^2 + C.$

17. $\int 3(\sin^{-1} x)^2/\sqrt{1-x^2} dx = (\sin^{-1} x)^3 + C.$

18. $\int x \cos (\sin^{-1} x) dx = -\frac{1}{3} (1-x^2)^{3/2} + C.$

19. $\int (1-x)/\sqrt{1-x^2} dx = \sin^{-1} x + \sqrt{1-x^2} + C.$

20. $\int (1-\tan^{-1} x)/(1+x^2) dx = \tan^{-1} x - \frac{1}{2} (\tan^{-1} x)^2 + C.$

21. Show that formulas I-8 and I-9 are not contradictory, since $\sin^{-1} x$ and $\cos^{-1} x$ differ by the constant $\pi/2$. Since an arbitrary constant belongs with each of these formulas, the constant $\pi/2$ can be absorbed by one of them.

22. Show that formulas I-10 and I-11 are not contradictory, as in Exercise 21.

13. The Differential Equation $y'' + a^2y = 0$. Exercises 21 and 22 of Art. 8 brought out an important fact concerning the functions $\sin x$ and $\cos x$. For each, the second derived function is the same as the original function, except for a change in sign. Thus,

$$D^2 \sin x = -\sin x, \quad \text{and} \quad D^2 \cos x = -\cos x,$$

or

$$D^2(\sin x) + \sin x = 0, \quad \text{and} \quad D^2(\cos x) + \cos x = 0.$$

Each of these functions is therefore a solution of the second order differential equation $D^2y + y = 0$, or $y'' + y = 0$ (see Chapter 6, Art. 10).

It can easily be verified that other solutions of this differential equation are $c_1 \sin x$, $c_2 \cos x$, and $c_1 \sin x + c_2 \cos x$, where c_1 and c_2 are any

constants whatever. The solution $y = c_1 \sin ax + c_2 \cos ax$ is general enough to include all possible solutions, although we shall not prove this.

In the same way, it can be shown that the differential equation

$$y'' + a^2y = 0 \quad (13.1)$$

has the solutions $c_1 \sin ax$, $c_2 \cos ax$, and the general solution $c_1 \sin ax + c_2 \cos ax$, where c_1 and c_2 are any constants. The constant a is not arbitrary, since it is given in the differential equation. This differential equation has numerous applications, especially in connection with oscillatory motion. Let us examine one of these.

We imagine a spring of natural length l (Fig. 24A). If a weight w is attached to the end of the spring it stretches to a length determined in

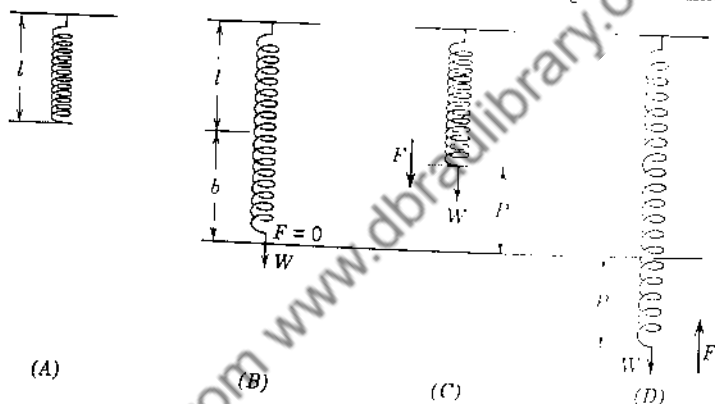


FIG. 24

accordance with Hooke's law. If this amount is b the law states that $w = kb$, where k is a constant for this particular spring, called the *spring constant*. The spring will remain in equilibrium at length $l + b$ if the weight remains attached and no further forces act (Fig. 24B). Now, suppose the weight is raised so that the spring takes the position shown in Fig. 24C and is then released. The weight will drop below the equilibrium point and then will oscillate above and below this position, because of the varying force in the spring and the momentum of the weight. At any time the force acting on the weight is the difference of two forces (if frictional forces are neglected), that due to gravity, i.e., the weight w acting downward, and that due to the tension in the spring acting upward.

If y represents the distance of the weight below the equilibrium point, positive if the weight is below and negative if the weight is above, the tension in the spring is $-k(y + b)$, negative because the tension is directed upward. Of course, if the weight is above the equilibrium point, y is

negative and $y + b$ is less than b . Since w is constant, the resultant force at any time is

$$\begin{aligned} F &= -k(y + b) + w = -ky - kb + w \\ &= -ky, \text{ since } w = kb. \end{aligned}$$

The weight thus oscillates up and down about its equilibrium position, subject to the force equal to $-ky$, which is proportional to its distance from the position of equilibrium and always directed toward that position (Fig. 25).

According to Newton's second law of motion, the acceleration of a body is proportional to the force acting on it; in fact, $a = F/m$, where m is the mass. The acceleration of the weight on the spring is therefore

$$a = -ky/m, \text{ or } a = -\alpha^2 y,$$

where α stands for the positive constant k/m .

Hence,

$$D_t^2 y = -\alpha^2 y, \text{ or } D_t^2 y + \alpha^2 y = 0. \quad (13.2)$$

This is the differential equation of the motion. It is a differential equation of the type discussed earlier in this article, and therefore has solutions of the form

$$y = c_1 \sin \alpha t, \quad y = c_2 \cos \alpha t, \quad \text{and} \quad y = c_1 \sin \alpha t + c_2 \cos \alpha t.$$

If we think of $t = 0$ as corresponding to an instant when the weight passes downward through the equilibrium position, we have $y = 0$ when $t = 0$. This condition is satisfied only by solutions of the form $y = c_1 \sin \alpha t$. The value of c_1 , which is the amplitude of the function $c \sin \alpha t$, represents the maximum displacement of the weight from its central position. In the example described, it is P (Fig. 27). Thus, $y = P \sin \alpha t$ is the equation of the oscillatory motion.

The general solution of the differential equation 13.2, which has the form

$$y = c_1 \sin \alpha t + c_2 \cos \alpha t, \quad (13.3)$$

can also be written

$$y = A \sin (\alpha t + B), \quad (13.4)$$

in which

$$A = \sqrt{c_1^2 + c_2^2} \text{ and } B = \cos^{-1}(c_1/\sqrt{c_1^2 + c_2^2}) = \sin^{-1}(c_2/\sqrt{c_1^2 + c_2^2}).$$

This solution represents an oscillatory motion of amplitude A and having $y = 0$ when $t = -B/\alpha$. Another form of the general solution is

$$y = A \cos (\alpha t + E), \quad (13.5)$$

in which $E = B - \pi/2$.

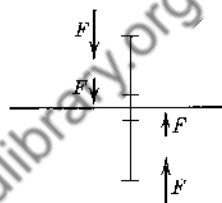


FIG. 25

The motion just discussed is an example of *simple harmonic motion*. This kind of motion is defined by any equation of the form (13.3), (13.4), or (13.5), or by a differential equation of the form (13.2). In Eq. 13.4 or 13.5, the constant A is the number of units in the amplitude, the period is $2\pi/\alpha$, and the fraction B/α determines the *phase* of the motion: its stage at a specified time, such as $t = 0$.

Simple harmonic motion is important as the basis of the study of oscillatory motion of all sorts. The study of such motion is fundamental for an understanding of sound, light, electromagnetics, and many other physical phenomena.

EXERCISES

1. Simple harmonic motion is often described as follows: if a point moves at a uniform speed around the circumference of a circle, the motion of its projection on any diameter of the circle is a simple harmonic motion. Show that this description is consistent with the definition of simple harmonic motion given in this article.
2. Write expressions for the velocity of the simple harmonic motions represented by Eqs. 13.3, 13.4, and 13.5.
3. Show that the simple harmonic motion in which $x = A$ (the amplitude) when $t = 0$ is represented by $x = A \cos \alpha t$.
4. At what points in a simple harmonic motion is the speed greatest? At what points is it least?
5. Show that the acceleration of a particle in simple harmonic motion is 0 when the particle is at its central position. When is the magnitude of the acceleration greatest?
6. Draw a graph of the simple harmonic motion represented by $y = 2 \sin \pi t$, plotting t on the horizontal axis and y on the vertical axis.
7. A point P describes simple harmonic motion in accordance with the differential equation $D_t^2 x + 40,000x = 0$, $x = 0$ when $t = 0$, and the maximum value of x is .5. Write the equation of the motion.
8. A point P describes simple harmonic motion on the interval of the x -axis, $-2 \leq x \leq 2$, starting at $x = 2$. If the unit of length is 1 in. and the unit of time is 1 sec, and the maximum acceleration is 8 in. per sec per sec, (a) write the differential equation of the motion, (b) the equation expressing x as a function of t , and (c) find the maximum velocity.
9. Write general solutions of the differential equations:
 (a) $D_x^2 y + x = 0$; (b) $D_t^2 s + ks = 0$.
10. Find a solution of the differential equation $D_t^2 y + 16y = 0$, if $y = 9$ and $D_t y = -48$ when $t = 0$.

14. Definite Integrals. Many definite integrals involving trigonometric functions can be found by making use of the formulas in Art. 12 and applying the fundamental theorem of calculus. The following examples illustrate this.

EXAMPLE 1. Calculate the area bounded by one arch of the curve $y = \sin x$ and the x -axis (Fig. 26).

SOLUTION. Applying the formula $A = \int_a^b f(x) dx$, we have

$$A = \int_0^{\pi} \sin x dx = \left[-\cos x \right]_0^{\pi} = 1 + 1 = 2.$$

EXAMPLE 2. Find the volume of revolution generated by revolving the area bounded by the curve $y = \tan x$, the x -axis, and the line $x = \pi/4$ about the x -axis.

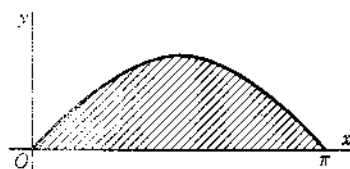


FIG. 26

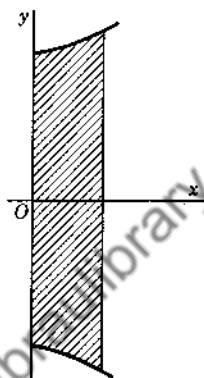


FIG. 27

SOLUTION. Applying formula 4.3 of Chapter 9,

$$V = \int_a^b \pi [f(x)]^2 dx$$

we have

$$\begin{aligned} V &= \int_0^{\pi/4} \pi \tan^2 x dx \\ &= \pi \int_0^{\pi/4} (\sec^2 x - 1) dx = \pi \int_0^{\pi/4} \sec^2 x dx - \pi \int_0^{\pi/4} 1 dx \\ &= \pi \left[\tan x \right]_0^{\pi/4} - \pi \left[x \right]_0^{\pi/4} = \pi(1 - \pi/4) = .674, \text{ approximately.} \end{aligned}$$

EXAMPLE 3. Calculate the area bounded by the curve $y^2(1 - x^2) = 1$, the y -axis, and the ordinate $x = \frac{1}{2}$ (Fig. 27).

SOLUTION. The curve $y^2(1 - x^2) = 1$ is symmetric about the x -axis, and hence the required area is equal to $2A$, where

$$\begin{aligned} A &= \int_0^{\frac{1}{2}} \frac{1}{\sqrt{1 - x^2}} dx \\ &= \left[\sin^{-1} x \right]_0^{\frac{1}{2}} = \sin^{-1} \frac{1}{2} - \sin^{-1} 0 = \pi/6. \end{aligned}$$

The whole area is therefore equal to $\pi/3$ units.

EXERCISES

Calculate the following definite integrals.

1. $\int_0^{\pi/2} \sin^3 x \cos x \, dx.$

2. $\int_{-\pi/6}^{\pi/6} \cos x \, dx.$

3. $\int_0^2 \frac{1}{4+x^2} \, dx.$

4. $\int_{-1}^0 \tan^2 x \, dx.$

5. $\int_{\pi/6}^{\pi/8} \tan x \sec^2 x \, dx.$

6. $\int_0^{1/\sqrt{2}} \frac{\sin^{-1} x}{\sqrt{1-x^2}} \, dx.$

7. $\int_0^{1/\sqrt{2}} \frac{x}{\sqrt{1-x^2}} \, dx.$

8. $\int_1^2 \frac{1}{x\sqrt{x^2-1}} \, dx.$

9. Find the area bounded by the curve $y = \sin x$, the y -axis, and the line $y = 1$ (Fig. 28).

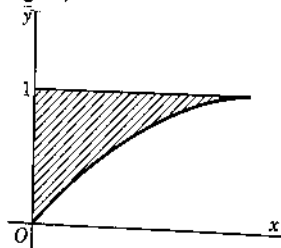


FIG. 28

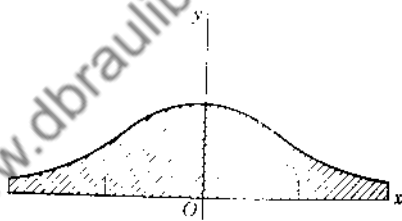


FIG. 29

10. Find the area bounded by the curve $y = \frac{1}{1+x^2}$, the x -axis, and the ordinates $x = -2$ and $x = 2$ (Fig. 29).

11. Find the volume of revolution generated by revolving about the x -axis the area bounded by the curve $y = \frac{1}{\sqrt{1+x^2}}$, the x -axis, and the ordinates $x = -1$ and $x = 1$.

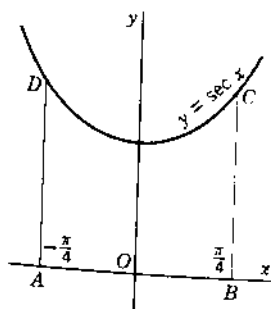


FIG. 30

12. Find the volume of revolution generated by revolving the area $ABCD$ in Fig. 30 about the x -axis.

Logarithmic and Exponential Functions

1. Introduction. We dealt with trigonometric functions in Chapter 10, starting with the familiar elementary definitions of the trigonometric ratios, and continuing with considerable dependence upon the geometric basis of these definitions. We shall follow a different procedure with the logarithmic function. There is a good reason for this. Although the elementary definition of a logarithm implies a logarithmic functional relation, $y = \log x$, it is only with difficulty that one can make use of this definition to establish the continuity of this function and differentiate it. We shall find that a more advanced definition will automatically dispose of the troublesome question of continuity. Moreover, the formula for the derivative of $\log x$ will follow directly from the definition of the function.

From the fundamental theorem in Chapter 8 we know that a function $\phi(x)$, determined by

$$\phi(x) = \int_a^x f(t) dt, \quad (1.1)$$

is continuous and has a derived function, and that

$$\phi'(x) = f(x).$$

In Chapter 8 we applied this result to the problem of integration, since it is an immediate corollary of this theorem that a definite integral can be found from a primitive function.

We shall now use Eq. 1.1 in a different way. Starting with a function $f(x)$, for which no primitive function is known, we shall define a new function $\phi(x)$ by means of (1.1). We shall then investigate the properties of this new function, with the help of the fundamental theorem. This is an example of a procedure that assumes great importance in advanced mathematics, viz., the discussion of a function by the examination of its properties. It is also noteworthy that the use of an integral to define a function is one of the most important applications of integration in advanced mathematics.

2. The Function $\int_1^x t^{-1} dt$. We have seen that inverse differentiation of any power of x always leads to the power one unit higher, with the single exception of x^{-1} . For this power we have no rule for inverse differentiation. Nevertheless, as long as the range of integration does not include the origin, the integral $\int_1^x t^{-1} dt$ exists and is a continuous differentiable function of x . This is guaranteed by the theorem on the existence of the integral of a continuous function and the fundamental theorem.

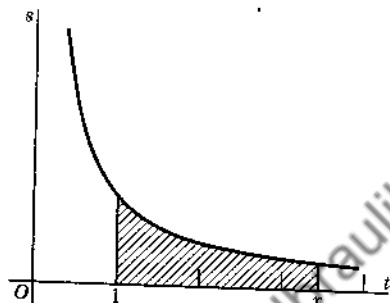


FIG. 1

Let us therefore investigate the function,

$$L(x) = \int_1^x t^{-1} dt = \int_1^x \frac{1}{t} dt, \quad \text{with } x > 0. \quad (2.1)$$

For any value of x that is greater than 1, $L(x)$ gives us the area bounded by the rectangular hyperbola $st = 1$, the t -axis, the ordinate $t = 1$, and the ordinate $t = x$ (Fig. 1). This is true, also, if $0 < x < 1$, although the area here has a negative sign. If $x = 1$, the area is 0, and so is

$L(1) = \int_1^1 t^{-1} dt$. From geometric considerations, it appears that $L(x)$ has the following properties.

- (1) $L(x)$ always increases with x , but at a continually decreasing rate.
- (2) $L(x)$ is negative for $x < 1$, equals 0 for $x = 1$, and is positive for $x > 1$.

From the definition of the function $L(x)$ by the integral in (1), we also know that

- (3) $L(x)$ is continuous for $x > 0$.
- (4) $L'(x) = 1/x$.

We must remember, too, that $L(x)$ is defined only for positive values of x .

The choice of 1 as the lower limit of the integral $\int_1^x \frac{1}{t} dt$ is arbitrary, and

a function $M(x) = \int_a^x \frac{1}{t} dt$, where both a and x are greater than 0, has properties (1), (3), and (4). In place of property (2) we would have

$M(x)$ negative if $x < a$, equal to 0 for $x = a$, and positive if $x > a$.

3. $L(x)$ and the Laws of Logarithms. The familiar laws of logarithms are the following.

$$\text{I. } \log ab = \log a + \log b.$$

$$\text{II. } \log a/b = \log a - \log b.$$

$$\text{III. } \log a^n = n \log a.$$

In these statements, a and b represent any positive numbers and n any rational number. We shall now show that the function $L(x)$ also has these properties. That is,

$$L(ab) = L(a) + L(b), \quad (3.1)$$

$$L(a/b) = L(a) - L(b), \quad (3.2)$$

$$L(a^n) = nL(a). \quad (3.3)$$

In order to prove (3.1), let us consider the function $L(ax)$. Differentiating, we have

$$D_x L(ax) = D_u L(u) \cdot D_x u, \text{ where } u = ax.$$

$$= \frac{1}{u} D_x u = \frac{1}{ax} a = \frac{1}{x}.$$

But it is true, also, that

$$D_x L(x) = 1/x.$$

Thus, since $L(x)$ and $L(ax)$ are two primitive functions of the same derived function $1/x$, they can differ only by a constant. Hence,

$$L(ax) = C + L(x).$$

If we now put $x = 1$, we have

$$L(a) = C + L(1) = C + 0.$$

Thus

$$L(ax) = L(a) + L(x).$$

Now, putting $x = b$,

$$L(ab) = L(a) + L(b), \quad (3.1)$$

which is the result to be established.

Suppose we let $a = 1/b$ in the result just proved. We then have

$$L(1/b \cdot b) = L(1) = L(1/b) + L(b).$$

Since $L(1) = 0$, it follows that

$$L(1/b) = -L(b).$$

Now, (3.1) holds for any positive values of a and b , and hence for a and $1/b$. That is,

$$L\left(\frac{a}{b}\right) = L\left(a \cdot \frac{1}{b}\right) = L(a) + L\left(\frac{1}{b}\right).$$

Since $L(1/b) = -L(b)$,

$$L(a/b) = L(a) - L(b). \quad (3.2)$$

The third property follows immediately from (3.1) if n is a positive integer. If $n = 4$, for example,

$$L(a^4) = L(a \cdot a \cdot a \cdot a) = L(a) + L(a) + L(a) + L(a) = 4L(a).$$

If n is a negative integer, we obtain (3.3) from (3.1) and (3.2). If $n = -3$, for example,

$$\begin{aligned} L(a^{-3}) &= L([1/a]^3) = L(1/a) + L(1/a) + L(1/a) \\ &= -L(a) - L(a) - L(a) = -3L(a). \end{aligned}$$

Now, suppose n is a fraction, for example, $\frac{1}{3}$. By (3.1),

$$\begin{aligned} L(a^{\frac{1}{3}}) + L(a^{\frac{1}{3}}) + L(a^{\frac{1}{3}}) &= L(a^{\frac{1}{3}} \cdot a^{\frac{1}{3}} \cdot a^{\frac{1}{3}}) \\ 3L(a^{\frac{1}{3}}) &= L(a), \end{aligned}$$

and hence,

$$L(a^{\frac{1}{3}}) = \frac{1}{3}L(a).$$

If n is a fraction like $\frac{2}{3}$, the result follows in the same way. Again

$$\begin{aligned} L(a^{\frac{2}{3}}) + L(a^{\frac{2}{3}}) + L(a^{\frac{2}{3}}) &= L(a^{\frac{2}{3}} \cdot a^{\frac{2}{3}} \cdot a^{\frac{2}{3}}) \\ 3L(a^{\frac{2}{3}}) &= L(a^2) = 2L(a). \\ L(a^{\frac{2}{3}}) &= \frac{2}{3}L(a). \end{aligned}$$

The proofs in the general cases, i.e., $n =$ any positive integer, any negative integer, or any fraction, are exactly like these illustrations.

Since the function $L(x)$ obeys all the rules of logarithms, it seems reasonable to call it a logarithmic function, and write

$$L(x) = \log x.$$

The common logarithms, having 10 as a base, are defined as follows.

$$\text{If } 10^L = N, \text{ then } \log_{10} N = L.$$

We know that any other positive number (except 1) could be used as a base, and all the laws of logarithms would still hold. Since the function $L(x)$ has been called a logarithmic function because of its having the properties of logarithms rather than by direct identification of its values with those obtained from a logarithmic table, one may naturally question

whether $L(x)$ is equal to $\log_b x$, where $\log_b x$ means the logarithm of x to a base b , defined in the elementary way. This will be shown in Art. 7. If this fact is granted, there is still the question as to the value of b . For example, is it true that $L(10) = \log_{10} 10 = 1$? The student can satisfy himself that this is not true, by considering the integral $\int_1^{10} \frac{1}{t} dt$ as an area. It is easily shown that $\int_1^{10} \frac{1}{t} dt > 1$. We shall determine b in the equation $L(x) = \log_b x$ in Art. 5.

EXERCISES

1. Obtain a lower approximating sum for the area represented by $L(10)$, by using rectangles one unit in width, and thus show that $L(10) > 1.9$. Similarly, by means of an upper approximating sum, show that $L(10) < 2.9$.

2. By a method similar to that of Exercise 1, using rectangles $\frac{1}{2}$ unit in width, show that $L(10)$ lies between 2.09 and 2.64.

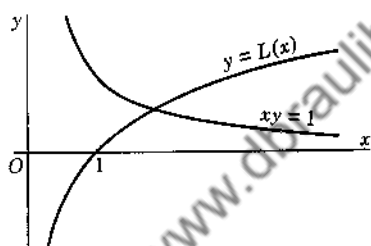


FIG. 2

4. Further Properties of $L(x)$. It was observed in Art. 2 that $L(x)$ is a continuous function that always increases with x , being negative for $x < 1$, equal to 0 for $x = 1$, and positive for $x > 1$. We can now show that $L(x)$ is large negatively for x near 0, and large positively for sufficiently large values of x . The last statement is easily seen as follows. If N is any large positive number, we consider $L(x)$ for $x > 10^N$. We have

$$L(x) > L(10^N) = N \cdot L(10) > N, \text{ since } L(10) > 1.$$

Hence, for any value of N it is possible to find a number 10^N such that when x surpasses this number $L(x) > N$. That is, $L(x) \rightarrow \infty$ as $x \rightarrow \infty$.

In the same way, if we put $x < \frac{1}{10^N}$,

$$L(x) < -N \cdot L(10) = -N,$$

and it follows that $L(x) \rightarrow -\infty$ as $x \rightarrow 0$.

The graph of $y = L(x)$ is shown in Fig. 2. The lengths of the ordinates of the curve can be estimated from the areas under the hyperbola $xy = 1$, and the slope is, of course, always equal to $1/x$.

5. The Base of the Logarithmic Function $L(x) = \log x$. The proof that $L(x)$ is identical with a logarithmic function $\log_b x$, for some base b , as defined in the elementary way, will be postponed to Art. 7, when we consider the properties of the inverse function to $L(x)$. In the present article we take this for granted, and make it our problem to determine b . For this purpose, we note that

$$L'(x) = 1/x, \quad L'(1) = 1.$$

But

$$L'(1) = \lim_{h \rightarrow 0} \frac{L(1+h) - L(1)}{h} = \lim_{h \rightarrow 0} 1/h \cdot L(1+h),$$

since

$$L(1) = 0.$$

Since, by (3.3),

$$1/h \cdot L(1+h) = L[(1+h)^{1/h}],$$

we have

$$L'(1) = 1 = \lim_{h \rightarrow 0} L[(1+h)^{1/h}]. \quad (5.1)$$

Now, from its definition as an integral, $L(x)$ is continuous. Hence,

$$\lim_{h \rightarrow 0} L(1+h)^{1/h} = L[\lim_{h \rightarrow 0} (1+h)^{1/h}].* \quad (5.2)$$

Putting this result in (5.1), we have

$$L[\lim_{h \rightarrow 0} (1+h)^{1/h}] = 1. \quad (5.3)$$

Now, if $(1+h)^{1/h}$ tends to a definite limit a , as h tends to 0, we have from (5.3), $L(a) = 1$. But if, as we have temporarily assumed, $L(x) = \log_b x$, then $\log_b a = 1$. From elementary theory of logarithms we know that the latter statement is true if, and only if, $b = a$. Therefore, the value of b is $\lim_{h \rightarrow 0} (1+h)^{1/h}$, if this limit exists.

In order to investigate $\lim_{h \rightarrow 0} (1+h)^{1/h}$ we shall calculate the values of $(1+h)^{1/h}$ for some of the terms of the sequence of values of h : 1, .1, .01, .001, $\dots$, which has the limit 0.

$$h = 1, \quad (1+h)^{1/h} = (1+1)^1 = 2^1 = 2.$$

$$h = .1, \quad (1+h)^{1/h} = (1+.1)^{10} = 1.1^{10} = 2.5938.$$

* If $f(x)$ is continuous at $x = a$, $\lim_{x \rightarrow a} f(x) = f(a) = f(\lim_{x \rightarrow a} x)$. If it turns out that there is a definite limit, $\lim_{h \rightarrow 0} (1+h)^{1/h}$, denoted by the letter e , for example, then, representing $(1+h)^{1/h}$ by u , statement (5.2) is the same as $\lim_{u \rightarrow e} L(u) = L(\lim_{u \rightarrow e} u)$.

$$h = .01, \quad (1 + h)^{1/h} = (1 + .01)^{100} = 1.01^{100} = 2.6992.$$

$$h = .001, \quad (1 + h)^{1/h} = (1 + .001)^{1000} = 1.001^{1000} = 2.7171.$$

$$h = .0001, \quad (1 + h)^{1/h} = (1 + .0001)^{10000} = 1.0001^{10000} = 2.718.$$

These results were obtained by logarithms and are accurate to the number of decimal places given. It appears from the sequence of values obtained that $(1 + h)^{1/h}$ increases as h becomes smaller, but that the increase is less and less rapid as x tends to 0, and that we might expect that there would be a limit not much greater than 2.718. Although we shall not prove this here,* it can be shown that there is, in fact, a limit, which is denoted by e , and that e is slightly greater than 2.718. There is no exact decimal representation of e with a limited number of decimal places, since this number is irrational.

It follows from the above that $b = e$. Thus, the logarithmic function

$$L(x) = \log x = \int_1^x \frac{1}{t} dt$$

has the base e ; that is, $L(x) = \log_e x$. Logarithms to the base e are called *natural logarithms*, or *Napierian logarithms*, because of their close relationship to the logarithms invented by Napier.† Because of the natural way in which they arise in calculus and the resulting simplicity in work depending upon calculus operations, natural logarithms are almost universally used in preference to all others in theoretical work. We shall use them in most of our future work. A table of natural logarithms should be readily available.‡ While it may be desirable on occasion to emphasize the fact that a logarithm has the base e , by the use of the subscript e — $\log_e x$, as a rule we shall write the natural logarithm without a subscript. If any other base is intended, a subscript will so indicate. Thus, $\log N$ means $\log_e N$, whereas $\log_{10} N$ will not be written without the subscript.

By means of the formula for change of base,

$$\log_a N = \frac{\log_b N}{\log_b a},$$

* See Courant, Vol. I, p. 43.

† The first table of logarithms was published by John Napier in 1614.

‡ E.g., *Mathematical Tables from Handbook of Chemistry and Physics*. Chemical Rubber Publishing Co., Cleveland.

the logarithmic function, $\log x$, can be expressed in terms of logarithms to other bases, and vice versa. Thus, for example,

$$\log x = \frac{\log_{10} x}{\log_{10} e},$$

and

$$\log_{10} x = \frac{\log x}{\log 10}.$$

By the use of different bases many logarithmic functions are possible, besides the one defined in Art. 2. They do not differ essentially, and all can easily be transformed into the natural logarithmic function.

EXERCISES

1. Letting $h = 1/n$, $(1 + h)^{1/h}$ becomes $(1 + 1/n)^n$ and $n \rightarrow \infty$ as $h \rightarrow 0$. Expand $1 + 1/n)^n$ by the binomial theorem, and hence show that $(1 + 1/n)^n < 1 + 1/1! + 1/2! + 1/3! + \dots + 1/n!$.
2. Show that $1 + 1/1! + 1/2! + 1/3! + \dots + 1/n! < 1 + 1 + 1/2 + 1/2^2 + \dots + 1/2^{n-1} < 3$, and hence that $(1 + 1/n)^n < 3$. From this, what can we conclude about the quantity, $\lim_{n \rightarrow \infty} (1 + 1/n)^n$, which is equal to $\lim_{h \rightarrow 0} (1 + h)^{1/h}$, assuming its existence?

6. Differentiation of Logarithmic Functions. From the definition of the natural logarithmic function

$$\log x = \int_1^x \frac{1}{t} dt,$$

it follows immediately that $D \log x = 1/x$. In this definition it is presumed that $x > 0$. Negative values can be brought within the compass of logarithmic functions by the use of absolute values. We see from

Fig. 3 that if $x < 0$, $\int_{-1}^x \frac{1}{t} dt = \int_1^{-x} \frac{1}{t} dt$. But

$$\int_1^{-x} \frac{1}{t} dt = \int_1^{|x|} \frac{1}{t} dt = \log |x|.$$

Hence,

$$\int_{-1}^x \frac{1}{t} dt = \log |x|.$$

It follows that

$$D \log |x| = 1/x. \quad (D-14)$$

This formula is valid for both positive and negative values of x , and fails only if $x = 0$, where $\log |x|$ is not defined. Generalizing, by the chain rule,

$$D_x \log |u| = 1/u D_x u. \quad (D-14a)$$

For logarithmic functions to bases other than e , we have the formulas

$$D \log_b |x| = \frac{1}{\log b} \cdot \frac{1}{x} \quad \text{and} \quad D_x \log_b |u| = \frac{1}{\log b} \cdot \frac{1}{u} D_x u.$$

These results follow from the fact that $\log_b |x| = \frac{\log |x|}{\log b}$ and $\log_b |u| = \frac{\log |u|}{\log b}$. They can be put in more useful form by replacing $\frac{1}{\log b}$ by $\log_b e$, which results from the change of base formula,

$$\log_b e = \frac{\log e}{\log b} = \frac{1}{\log b}.$$

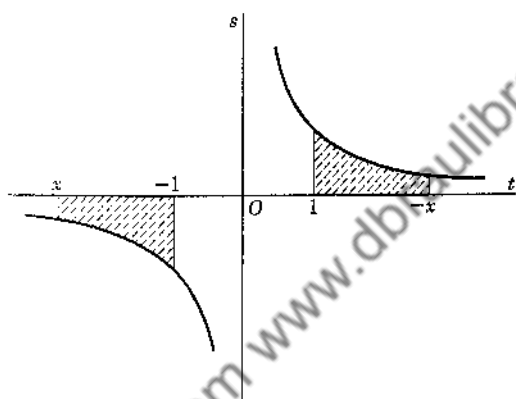


FIG. 3

Thus

$$D \log_b |x| = \log_b e \cdot \frac{1}{x} \quad (D-15)$$

and

$$D_x \log_b |u| = \log_b e \cdot \frac{1}{u} D_x u. \quad (D-15a)$$

EXAMPLE 1. Find Dy if $y = \log(x^2 + 1)$.

SOLUTION. Using formula $D-14a$,

$$\begin{aligned} Dy &= D \log(x^2 + 1) \\ &= \frac{1}{x^2 + 1} \cdot 2x = \frac{2x}{x^2 + 1} \end{aligned}$$

EXAMPLE 2. Find Dy if $y = \log_{10} \sqrt{x^2 + 1}$.

SOLUTION. We first write $\log_{10} \sqrt{x^2 + 1} = \frac{1}{2} \log_{10} (x^2 + 1)$. Now, using formula D-15a,

$$\begin{aligned} Dy &= D\left[\frac{1}{2} \log_{10} (x^2 + 1)\right] \\ &= \frac{1}{2} \cdot \frac{1}{x^2 + 1} \cdot 2x \cdot \log_{10} e \\ &= \frac{x}{x^2 + 1} \cdot \log_{10} e. \end{aligned}$$

EXAMPLE 3. Find Dy if $y = \log \left| \frac{1 + x^2}{1 - x^2} \right|$.

SOLUTION. We first write $\log \left| \frac{1 + x^2}{1 - x^2} \right| = \log |1 + x^2| - \log |1 - x^2|$.

Differentiating, we find

$$Dy = \frac{2x}{1 + x^2} + \frac{2x}{1 - x^2} = \frac{4x}{1 - x^4}.$$

EXERCISES

Verify the differentiations in Exercises 1-20 and state the range of validity of the result.

1. $D \log (x + 2) = 1/(x + 2)$.
2. $D \log |x^3| = 3/x$.
3. $D \log (1/x) = -1/x$.
4. $D \log |(x^2 + 1)/(x + 1)| = 2x/(x^2 + 1) - 1/(x + 1)$.
5. $D(x \log x) = 1 + \log x$.
6. $D \log |\sin x| = \cot x$.
7. $D \log \sec x = \tan x$.
8. $D[(x^2 + 1) \log (x^2 + 1)] = 2x[1 + \log (x^2 + 1)]$.
9. $D(\log x)^3 = 3(\log x)^2/x$.
10. $D \log(\log x) = 1/(x \log x)$.
11. $D \log_{10} x \sqrt{1 - x^2} = [1/x - x/(1 - x^2)] \log_{10} e$.
12. $D \log_{10} |\tan x| = (\tan x + \cot x) \log_{10} e$.
13. $D \log \sqrt{(1 - x)/(1 + x)} = -1/(1 - x^2)$.
14. $D \log_{10} (x^2 + x) = [1/x + 1/(x + 1)] \log_{10} e$.
15. $D \log (x + \sqrt{1 + x^2}) = 1/\sqrt{1 + x^2}$.
16. $D \log \sqrt{(1 - \cos x)/(1 + \cos x)} = \csc x$.
17. $D \log |\tan (\pi/4 + x/2)| = \sec x$.
18. $D \log \sin^2 x = 2 \cot x$.

19. $D \log [x^2 \sqrt{1+x^2}] = (4x^2 + 3)/[x(1+x^2)]$.
20. $D \log |\sin^{-1} x| = 1/(\sqrt{1-x^2} \sin^{-1} x)$.
21. Find the slope of the graph of $y = \log_{10} x$ at $x = 1$; $x = \frac{1}{2}$; $x = \frac{1}{10}$; $x = 10$.
22. Find the slope of the graph of $y = \log x$ at $x = 1$; $x = .1$; $x = 10$.
23. Find the minimum point of $y = (\log x)^2$, and the point of inflection. Draw the graph.
24. Show that the differential equation, $x^2 D^2 y - x Dy + y = 0$, is satisfied by $y = Ax \log x + Bx$, where A and B are any constants.
25. Show that the differential equation, $x^2 D^2 y + x Dy + y = 0$, is satisfied by $y = A \sin \log x + B \cos \log x$, where A and B are any constants.

7. The Inverse of $L(x)$. In this article we do *not* assume as in Arts. 4-6, that $L(x)$ is identical with $\log_b x$ for some base b , and the consequent result in Art. 5, $L(x) = \log_e x$. We have yet to justify this identity, and we shall do so now.

Just as we formed inverse functions of the trigonometric functions, so we can form the inverse function of $L(x)$. We observe that $L(x)$ is a monotonic increasing function throughout its range. Hence, we do not have to restrict the range in which we can form the inverse function as we did for the inverse trigonometric functions. We write

$$y = L(x), \quad x = E(y)$$

as inverse ways of expressing the same functional relation. From our knowledge of $L(x)$, we know that x is always positive whereas y takes on all values from $-\infty$ to ∞ . Hence, in the second equation, also, the range of y includes all real values, whereas x takes on positive values only.

As usual when studying inverse relations, we shall interchange the variables in the second equation so that x is the independent variable, and study the function

$$y = E(x).$$

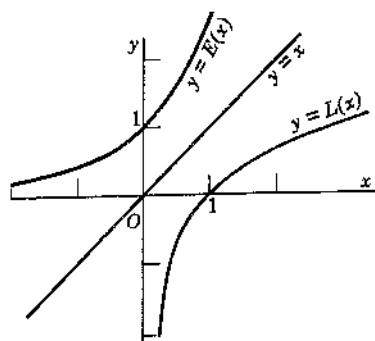


FIG. 4

Figure 4 shows the graph of this function, which can be obtained from that of $y = L(x)$ by reflection in the line $y = x$. From the graph we see that $E(x)$ has the following properties. (1) $E(x)$ is defined for all values of x and is always positive. (2) $E(0) = 1$. (3) $E(x)$ tends to 0 as x tends to $-\infty$. (4) $E(x)$ becomes very great as x increases. (5) $E(x)$ is continuous for all values of x .

From (3.1), (3.2), and (3.3) that apply to $L(x)$ we can derive corresponding rules of combination for the inverse function $E(x)$. First, let us consider $E(x)$ for two values of x , say a and b . Suppose

$$E(a) = \alpha, \quad E(b) = \beta.$$

In terms of the function $L(x)$, these statements become

$$a = L(\alpha), \quad b = L(\beta).$$

By (3.1),

$$a + b = L(\alpha) + L(\beta) = L(\alpha\beta).$$

Expressing this relation by the inverse function $E(x)$, we have

$$\alpha\beta = E(a + b).$$

Hence,

$$E(a) \cdot E(b) = E(a + b). \quad (7.1)$$

This multiplication theorem reminds one of the rule for multiplying powers of the same base, i.e.,

$$p^a \cdot p^b = p^{a+b}.$$

It therefore seems possible that $E(x)$ may be a function of the form p^x , called an *exponential* function. We shall show that this is true.

We observe at the outset that $E(0) = 1$, as would be true if $E(x)$ were an exponential function, since $p^0 = 1$ for any p . Second, we note that since $L(x)$ is continuous and increases from 0 to ∞ as x increases from 1 to ∞ , it must take on the value 1 for some value of x greater than 1. This number we denote by p and write $L(p) = 1$. Consequently,

$$E(1) = p. \quad (7.2)$$

Using (7.1), we have

$$E(1) \cdot E(1) \cdot \dots \text{(to } n \text{ factors)} = E(n).$$

Since $E(1) = p$, this is equivalent to

$$E(n) = p^n.$$

Again, $[E(1/n)]^n = E(1/n) \cdot E(1/n) \cdot \dots \text{(to } n \text{ factors)} = E(n \cdot 1/n)$, by (7.1). That is,

$$[E(1/n)]^n = E(1) = p, \text{ by (7.2).}$$

Hence,

$$E(1/n) = p^{1/n}. \quad (7.3)$$

Now, $E(1/n) \cdot E(1/n) \cdot \dots \text{(to } m \text{ factors)} = E(m \cdot 1/n)$, by (7.1), and, by (7.3), this product is equal to $p^{m/n}$. Therefore,

$$E(m/n) = p^{m/n}. \quad (7.4)$$

Thus, $E(x) = p^x$ for all positive rational values of x and for $x = 0$. To show that this equality holds for negative rational values of x we note that, by (7.1),

$$E(x) \cdot E(-x) = E(x - x) = E(0) = 1,$$

whence

$$E(-x) = 1/E(x).$$

Thus, if x is a positive rational number, and consequently $-x$ is a negative rational number,

$$E(-x) = 1/p^x = p^{-x}.$$

We have therefore identified the function $E(x)$ with p^x for all rational values of x .

Since $E(x)$ is a continuous function, its value for any irrational value of x , say $\bar{x}$, is equal to the limit of $E(x)$ as x approaches $\bar{x}$. In this approach, x can take on a sequence of rational values, for which it has been shown that $E(x) = p^x$. Since $p^{\bar{x}}$ is defined in the same way, i.e., the limit of p^x as x approaches $\bar{x}$, we are assured that $E(x) = p^x$ for all real values of x .*

This result settles the matter left open in Arts. 3 and 5, as to whether the function $L(x)$ could be thought of as a logarithmic function in the sense of the elementary definition of a logarithm. Going back to the beginning of this article, we write in place of $x = E(y)$,

$$x = p^y.$$

Hence, by the elementary definition,

$$y = \log_p x.$$

But, since $x = E(y)$, by the definition of $E(y)$, $y = L(x)$. Hence,

$$L(x) = \log_p x.$$

It will be remembered that it was shown in Art. 5 that if $L(x)$ were identical with some logarithmic function $\log_b x$, defined in the elementary way, then $b = e$. Hence, in the above discussion of $E(x)$, we can replace p by e and write

$$E(x) = e^x.$$

Henceforth we shall dispense with the notations $L(x)$ and $E(x)$, replacing them by $\log x$ and e^x , respectively.

* Actually, p^x is usually defined only for rational values in elementary books. For a complete discussion see Knopp, *Theory and Application of Infinite Series*, pp. 47 ff. Blackie, London, 1928.

EXERCISES

1. Simplify without rewriting in logarithmic and exponential form:

$$(a) L[E(a)].$$

$$(b) E[L(b)].$$

$$(c) L[E(a) \cdot E(b)].$$

$$(d) L[\sqrt{E(a)}].$$

$$(e) E[L(a) + L(b)].$$

$$(f) E[nL(a)].$$

$$(g) E[8L(a)] - E[2L(a) + 3L(b)].$$

$$(h) L\{E(1/x)^2\}.$$

2. Rewrite each of the expressions in Exercise 1, using logarithmic and exponential notation, and thus check the results of Exercise 1.

3. From $E(a) = \alpha$, we have $L(\alpha) = a$. Hence, $L(\alpha^n) = na$. From this show that $[E(a)]^n = E(na)$.

8. Other Exponential Functions. Occasionally, we may encounter such exponential functions as 2^x , 10^x , or b^x , where b is any positive number other than 1 (if $b = 1$ the function is constant). Such functions can be transformed without difficulty into exponential functions with base e . This is done by means of the fundamental identity

$$a = e^{\log_e a},$$

Thus,

$$2^x = (e^{\log_e 2})^x = e^{\log_e 2 \cdot x},$$

$$10^x = (e^{\log_e 10})^x = e^{\log_e 10 \cdot x},$$

$$b^x = (e^{\log_e b})^x = e^{\log_e b \cdot x}.$$

9. Differentiation of Exponential Functions. The derivative of the exponential function e^x is easily obtained from that of $\log x$, by means of the rule for differentiating inverse functions. Thus, if

$$y = e^x, \quad x = \log_e y.$$

$$D_x y = \frac{1}{D_y x} = \frac{1}{1/y} = y = e^x.$$

Hence,

$$De^x = e^x. \quad (D-16)$$

Generalizing, by the chain rule,

$$D_x e^u = e^u D_x u. \quad (D-16a)$$

Exponential functions with other bases can be differentiated by changing them to the form e^u , as in Art. 8, and applying (D-16a). Thus, since

$$b^x = e^{\log_e b \cdot x} = e^u, \text{ where } u = \log_e b \cdot x,$$

$$D_x b^x = e^{\log_e b \cdot x} \log_e b = b^x \log_e b. \quad (D-17)$$

Similarly,

$$D_x b^u = b^u \log_e b D_x u. \quad (D-17a)$$

The graph of the function $y = e^x$ has the property that the slope of the tangent at any point is equal to the ordinate of the point. In the language of rates, the rate of change of the function e^x at any value of x is numerically equal to the functional value itself. For example, the slope of the graph of $y = e^x$ at $x = 0, y = 1$, is 1, and thus the function is changing at the rate of one unit per unit change of x when the function also has the value of one unit. Similarly, the graph of the function $y = b^x$ has the property that the slope of the tangent at any point is proportional to the ordinate of the point, the proportionality factor being $\log_e b$. It is the fact that an exponential function always changes at a rate that is proportional to its own value which makes these functions so useful in the representation of growth and decay. This application will be discussed more fully in Art. 15.

EXAMPLE 1. Find $f'(x)$ if $f(x) = e^{-x^2}$.

SOLUTION. Using formula D-16a,

$$\begin{aligned} f'(x) &= e^{-x^2} D(-x^2) \\ &= -2xe^{-x^2}. \end{aligned}$$

EXAMPLE 2. Find Dy and D^2y if $y = \frac{e^x - e^{-x}}{2}$.

SOLUTION. Using formulas D-16 and D-16a,

$$\begin{aligned} Dy &= \frac{1}{2}(e^x + e^{-x}), \\ D^2y &= \frac{1}{2}(e^x - e^{-x}) = y. \end{aligned}$$

This result for the second derivative reminds one of the second derivative of $y = \sin x$, namely, $D^2y = -\sin x = -y$, and suggests a possible similarity between the functions $\frac{1}{2}(e^x - e^{-x})$ and $\sin x$ (see Chapter 18).

EXAMPLE 3. Find Dy if $y = e^{2x}(3 \sin 3x + 2 \cos 3x)$.

SOLUTION. Using the rules for differentiating sums and products, with formulas D-2a, D-3a, and D-16a,

$$\begin{aligned} Dy &= e^{2x}(9 \cos 3x - 6 \sin 3x) + 2e^{2x}(3 \sin 3x + 2 \cos 3x) \\ &= 13e^{2x} \cos 3x. \end{aligned}$$

EXAMPLE 4. Find $D(x^n 10^x)$.

SOLUTION. Using the rule for differentiating a product, with formulas D-1 and D-17,

$$\begin{aligned} D(x^n 10^x) &= x^n 10^x \log 10 + nx^{n-1} 10^x \\ &= x^{n-1} 10^x (x \log 10 + n). \end{aligned}$$

EXAMPLE 5. Find $f'(x)$ if $f(x) = e^{\tan^2 x}$.

SOLUTION. Using formula D-16a with (D-4) and the power rule,

$$f'(x) = e^{\tan^2 x} \cdot 2 \tan x \sec^2 x.$$

EXERCISES

Verify the differentiations in Exercises 1–24.

1. $De^{-x} = -e^{-x}$.
2. $Dce^{a^2x} = ca^2e^{a^2x}$.
3. $D2e^{-.1x} = -.2e^{-.1x}$.
4. $D.1e^{x^2} = .2xe^{x^2}$.
5. $D10^{x^2} = 2x \cdot 10^{x^2} \cdot \log 10$.
6. $D2(e^{x/2} + e^{-x/2}) = e^{x/2} - e^{-x/2}$.
7. $De^{2\theta}/4 = e^{2\theta}/2$.
8. $D2e^{\theta/4} = e^{\theta/4}/2$.
9. $De^{\sin t} = (\cos t)e^{\sin t}$.
10. $Dte^t = e^t(1 + t)$.
11. $De^x \sin x = e^x(\cos x + \sin x)$.
12. $De^{-t} \cos t = -e^{-t}(\cos t + \sin t)$.
13. $De^{1/x} = -e^{1/x}/x^2$.
14. $D \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{4}{(e^x + e^{-x})^2}$.
15. $D \sin e^x = e^x \cos e^x$.
16. $Dx^2b^{x^2} = x^2b^{x^2}(2/x + 2x \log b)$.
17. $D \log 2^{3x} = 3 \log 2$.
18. $D \log e^{x^2} = 2x$.
19. $D \sin^{-1} e^{-x} = -e^{-x}/\sqrt{1 - e^{-2x}}$.
20. $D \tan^{-1} e^{1x} = .1e^{1x}/(1 + e^{2x})$.
21. $D \log[e^x/(e^x - 1)] = -1(e^x - 1)$.
22. $D(e^{t/2} + e^{-t/2})^2 = e^t - e^{-t}$.
23. $De^x/x = e^x(1/x - 1/x^2)$.
24. $Dze^{1/z} = e^{1/z}(1 - 1/z)$.
25. Find the slope of the graph of $y = xe^x$ at $x = 0$, $x = 1$, $x = -1$, $x = -2$, and draw the curve.
26. Show that the graph of $y = xe^{-x}$ has a maximum point at $x = 1$, and draw the graph.
27. Find the level points and inflection points, if any, and draw the graph of $y = e^{-.1x^2}$.
28. Show that $y = c(e^x + e^{-x})$ satisfies the differential equation $D_x^2y - y = 0$.
29. Show that $y = e^x(A \cos x + B \sin x)$ satisfies the differential equation $D_x^2y - 2D_xy + 2y = 0$.
30. Show that $y = Ae^x + Bxe^x$ satisfies the differential equation $D_x^2y - 2D_xy + y = 0$.

10. A Logarithmic Simplification. The rule for differentiating implicit functions, first discussed in Chapter 3, Art. 13, can often be applied advantageously in conjunction with formulas for differentiating logarithmic functions. This is brought out in the following examples.

EXAMPLE 1. Find $D[\sqrt[3]{(x-1)^2/(x^2+1)}]$.

SOLUTION. We can first set $y = \sqrt[3]{(x-1)^2/(x^2+1)}$. If we now take the logarithms of the absolute values of both sides of this equation, we obtain the implicit relation

$$\log |y| = \log |\sqrt[3]{(x-1)^2/(x^2+1)}|,$$

which reduces to

$$\log |y| = \frac{2}{3} \log |x-1| - \log |x^2+1|.$$

Now, differentiating both sides with respect to x ,

$$(D_x y)/y = \frac{2}{3}[1/(x-1)] - 2x/(x^2+1).$$

Multiplying by y ,

$$\begin{aligned} D_x y &= \left\{ \frac{2}{3} [1/(x-1)] - 2x/(x^2+1) \right\} y \\ &= \left\{ \frac{2}{3} [1/(x-1)] - 2x/(x^2+1) \right\} \sqrt[3]{(x-1)^2/(x^2+1)}. \end{aligned}$$

EXAMPLE 2. Find $D_x y$ if $y = (\sin x)^x$.

SOLUTION. Taking the logarithms of the absolute values of both sides,

$$\log |y| = \log |(\sin x)^x| = x \log |\sin x|.$$

Differentiating,

$$(D_x y)/y = \log |\sin x| + x \cot x.$$

Multiplying by y ,

$$\begin{aligned} D_x y &= y(\log |\sin x| + x \cot x) \\ &= (\sin x)^x (\log |\sin x| + x \cot x). \end{aligned}$$

The method of these examples can be expressed by means of a single formula. Thus, since $(D_x u)/u = D(\log |u|)$, we have

$$Du = uD(\log |u|). \quad (10.1)$$

By means of formula 10.1 the solution of Example 1 becomes

$$\begin{aligned} D[\sqrt[3]{(x-1)^2/(x^2+1)}] &= [\sqrt[3]{(x-1)^2/(x^2+1)}] \cdot D \log |\sqrt[3]{(x-1)^2/(x^2+1)}| \\ &= [\sqrt[3]{(x-1)^2/(x^2+1)}] \cdot D[\frac{2}{3} \log |x-1| - \log |x^2+1|] \\ &= [\sqrt[3]{(x-1)^2/(x^2+1)}] \cdot [\frac{2}{3} \cdot 1/(x-1) - 2x/(x^2+1)]. \end{aligned}$$

Applying the method of the examples, or formula 10.1, we can derive a further general formula for differentiation. Consider the function u^v , in which both u and v are functions of x . By formula 10.1,

$$\begin{aligned} D_x(u^v) &= u^v D_x \log |u^v| \\ &= u^v D_x(v \log |u|) \\ &= u^v [v \cdot (D_x u)/u + (\log |u|) \cdot D_x v] \\ &= v u^{v-1} D_x u + u^v (D_x v) \log |u|. \end{aligned} \quad (10.2)$$

In Chapter 3 it was shown that $Du^n = nu^{n-1}$ for all rational values of n . By means of (10.2) we can now further extend this power formula to all real values of n . Since (10.2) holds for $v =$ any function that has a derivative, it holds for $v = n =$ any real number, with $D_x v = 0$. Thus,

$$\begin{aligned} D_x(u^n) &= nu^{n-1} + u^n \log |u| \cdot 0 \\ &= nu^{n-1}. \end{aligned}$$

EXERCISES

Use the method of Examples 1 and 2 or formula 10.1 to find $D_x y$ from each of the following.

1. $y = (x^2 + 1)/(x^2 - 1).$

2. $y = \sqrt{x-1}/(x^2 + 2).$

3. $y = (x^2 + 1)\sqrt{x^2 - 1}/(2x + 1).$

4. $y = \sqrt{(x-4)(x-3)}/(x-2)(x-1).$

5. $y = (\sin^2 x) \sqrt{(x+1)^3}.$

6. $y = xe^x \sin x.$

7. $y = x^x.$

8. $y = x^{(e^x)}.$

9. $y = e^{\sin x}.$

10. $y = \sqrt[3]{3x-1} \sqrt{(2x+1)^3}/[(x^2+4)(\sqrt[3]{x^3-8})].$

11. The Inverse Derivative of $1/x$. The statement that

$$D^{-1}1/x = \log |x| + C,$$

or

$$\int 1/x \, dx = \log |x| + C, \quad (I-14)$$

follows immediately from formula $D-14$. The more general formula is

$$D_x^{-1}[1/u \, D_x u] = \log |u| + C,$$

or

$$\int 1/u \, D_x u \, dx = \log |u| + C. \quad (I-14a)$$

Formula $I-14$ fills the sole remaining gap left by the formula for a primitive function of a power of x , and it is now possible to find $D^{-1}x^n$ for all values of n .

$$D^{-1}x^n = x^{n+1} + C, \text{ if } n \neq -1,$$

$$= \log |x| + C, \text{ if } n = -1.$$

Formula $I-14$ also calls attention to the fact that the transcendental function, $\log |x|$, is obtained from an algebraic function, $1/x$, by inverse differentiation. This kind of relationship between a transcendental function on the one hand and an algebraic function on the other was pointed out in connection with the inverse trigonometric functions.

By means of $(I-14a)$, some further inverse derivative formulas involving trigonometric functions can be obtained. Since $D \sin x = \cos x$,

$$D^{-1} \cot x = D^{-1} \frac{\cos x}{\sin x} = D^{-1} \left[\frac{1}{\sin x} D \sin x \right]$$

$$= D_x^{-1}[1/u \, D_x u],$$

where $u = \sin x$. Hence, by formula $I-14a$,

$$D^{-1} \cot x = \log |\sin x| + C,$$

or

$$\int \cot x \, dx = \log |\sin x| + C. \quad (I-18)$$

Similarly, since $D \cos x = -\sin x$,

$$\begin{aligned} D^{-1} \tan x &= D^{-1} \frac{\sin x}{\cos x} = -D^{-1} \left[\frac{1}{\cos x} D \cos x \right] \\ &= -D_x^{-1} [1/u \, D_x u], \end{aligned}$$

where $u = \cos x$. Hence,

$$D^{-1} \tan x = -\log |\cos x| + C = \log |\sec x| + C,$$

or

$$\int \tan x \, dx = \log |\sec x| + C. \quad (I-19)$$

By observing that $\sec x \tan x + \sec^2 x = D \sec x + D \tan x$ we see that

$$\sec x (\sec x + \tan x) = D(\sec x + \tan x)$$

and

$$\sec x = \frac{1}{\sec x + \tan x} D(\sec x + \tan x).$$

Hence

$$D^{-1} \sec x = D_x^{-1} [1/u \, D_x u] = \log |u| + C,$$

where

$$u = \sec x + \tan x.$$

That is,

$$D^{-1} \sec x = \log |\sec x + \tan x| + C,$$

or

$$\int \sec x \, dx = \log |\sec x + \tan x| + C. \quad (I-20)$$

In the same way, we observe that

$$-\csc x \cot x + \csc^2 x = D(\csc x - \cot x).$$

The student can now carry out steps similar to those leading to formula I-20 and obtain

$$D^{-1} \csc x = \log |\csc x - \cot x| + C,$$

or

$$\int \csc x \, dx = \log |\csc x - \cot x| + C. \quad (I-21)$$

Formulas I-2, I-3, I-18, I-19, I-20, and I-21 make it possible to write the primitive functions of all the six fundamental trigonometric functions at sight. Just as for every inverse derivative formula previously obtained, each of the formulas in this article can be extended by replacing x by a

function of x , provided that the derivative of this function is present as a factor. Thus, we have

$$\begin{aligned} \text{or } D_x^{-1}[\csc u D_x u] &= \log |\sin u| + C, \\ \int \csc u D_x u dx &= \log |\sin u| + C, \end{aligned} \quad (I-18a)$$

$$\begin{aligned} \text{or } D_x^{-1}[\tan u D_x u] &= \log |\sec u| + C, \\ \int \tan u D_x u dx &= \log |\sec u| + C. \end{aligned} \quad (I-19a)$$

$$\begin{aligned} \text{or } D_x^{-1}[\sec u D_x u] &= \log |\sec u + \tan u| + C, \\ \int \sec u D_x u dx &= \log |\sec u + \tan u| + C. \end{aligned} \quad (I-20a)$$

$$\begin{aligned} \text{or } D_x^{-1}[\csc u D_x u] &= \log |\csc u - \cot u| + C, \\ \int \csc u D_x u dx &= \log |\csc u - \cot u| + C. \end{aligned} \quad (I-21a)$$

EXERCISES

1. Complete the derivation of formula I-21.
2. Show that $\csc x - \cot x = \tan x/2$, and hence that $D^{-1} \csc x = \log |\tan x/2| + C$.
3. Show that $\sec x + \tan x = \tan(\pi/4 + x/2)$, and hence that $D^{-1} \sec x = \log |\tan(\pi/4 + x/2)| + C$.

Verify the primitive functions in Exercises 4-30.

4. $D^{-1} 5/x = 5 \log |x| + C$.
 5. $D^{-1} 1/2x = \frac{1}{2} \log |x| + C$.
 6. $D^{-1} 2x/(x^2 + 1) = \log(x^2 + 1) + C$.
 7. $D^{-1} 3x^2/(x^3 + 27) = \log |x^3 + 27| + C$.
 8. $D^{-1} (2x + 1)/(x^3 + x) = \log |x^2 + x| + C$.
 9. $D^{-1} 2x/(1 - x^2) = -\log |1 - x^2| + C$.
 10. $D^{-1} x/(4 + x^2) = \log \sqrt{4 + x^2} + C$.
 11. $D^{-1} (x + 1)/(x^3 + 2x + 4) = \log \sqrt{x^2 + 2x + 4} + C$.
 12. $D^{-1} (3x^2 + 2x + 1)/(x^3 + x^2 + x + 5) = \log |x^3 + x^2 + x + 5| + C$.
 13. $D^{-1} (x + 1)/(x - 1) = x + 2 \log |x - 1| + C$.
- [Hint: Put $(x + 1)/(x - 1) = 1 + 2/(x - 1)$.]
14. $D^{-1} (x^2 + x + 2)/(x^2 - x + 3) = x + \log |x^2 - x + 3| + C$.
 15. $D^{-1} 1/[2\sqrt{x}(\sqrt{x} + 1)] = \log(\sqrt{x} + 1) + C$.
 16. $\int (\log x)/x dx = \frac{1}{2} (\log x)^2 + C$.

$$17. \int 1/(x \log |x|) dx = \log |(\log |x|)| + C.$$

$$18. \int (\log |x|)^2/x dx = \frac{1}{3} (\log |x|)^3 + C.$$

$$19. \int 1/[x(1 + \log |x|)] dx = \log |(1 + \log |x|)| + C.$$

$$20. \int (\log |x| + 1)/[x(\log |x| - 1)] dx = \log |x| + 2 \log |(\log |x| - 1)|.$$

$$21. \int (\log \sqrt{x})/x dx = \frac{1}{4} (\log |x|)^2 + C.$$

$$22. \int e^x/(e^x - 1) dx = \log |e^x - 1| + C.$$

$$23. \int \tan 2x dx = -\frac{1}{2} \log |\sec 2x| + C.$$

$$24. \int \sec (x/2) dx = 2 \log |\sec (x/2) + \tan (x/2)| + C.$$

$$25. \int \sec^2 x / \tan x dx = \log |\tan x| + C.$$

$$26. \int \tan^3 x dx = \frac{1}{2} \tan^2 x + \log |\sec x| + C.$$

[Hint: Write $\tan^3 x = \tan x \sec^2 x - \tan x$.]

$$27. \int \cos x/(1 + \sin x) dx = \log (1 + \sin x) + C.$$

$$28. \int \csc x/[\tan x(1 + \csc x)] dx = -\log |1 + \csc x| + C.$$

$$29. \int 1/[(1 + x^2) \tan^{-1} x] dx = \log |\tan^{-1} x| + C.$$

$$30. \int 1/[\sqrt{a^2 - x^2} \sin^{-1} x] dx = \log |\sin^{-1} x| + C.$$

12. Inverse Derivatives of Exponential Functions. A group of inverse derivative formulas are obtained by inverting the formulas for differentiating the exponential functions. Thus,

$$D^{-1}e^x = e^x + C, \quad \text{or} \quad \int e^x dx = e^x + C. \quad (I-16)$$

Similarly,

$$D^{-1}b^x = b^x/\log b + C, \quad \text{or} \quad \int b^x dx = b^x/\log b + C. \quad (I-17)$$

These formulas are easily proved by differentiation, as are the more general formulas,

$$D_x^{-1}[e^u D_x u] = e^u + C, \quad \text{or} \quad \int e^u D_x u dx = e^u + C, \quad (I-16a)$$

$$D_x^{-1}[b^u D_x u] = b^u/\log b + C,$$

or

$$\int b^u D_x u dx = b^u/\log b + C. \quad (I-17a)$$

Formulas I-17 and I-17a can also be written

$$\int b^x dx = b^x \log_b e + C, \quad (I-17)$$

$$\int b^u D_x u dx = b^u \log_b e + C. \quad (I-17a)$$

Among the above formulas, (I-16a) is most often needed in applications, particularly with $u = kx$. In this form, it is easy to show that

$$D^{-1}e^{kx} = e^{kx}/k + C, \quad \text{or} \quad \int e^{kx} dx = e^{kx}/k + C.$$

EXAMPLE 1. Find $D^{-1}(e^{\sin 2x} \cos 2x)$.

SOLUTION. Since $2 \cos 2x = D \sin 2x$, we have to evaluate

$$\bullet \quad D^{-1}e^{\sin 2x} \left(\frac{1}{2} D \sin 2x\right) = \frac{1}{2} D_x^{-1}[e^u D_x u],$$

where $u = \sin 2x$. Hence,

$$D^{-1}(e^{\sin 2x} \cos 2x) = \frac{1}{2} e^{\sin 2x} + C.$$

EXAMPLE 2. Find $\int x 10^{-x^2} dx$.

SOLUTION. Since $D(-x^2) = -2x$, we have to evaluate

$$\int 10^{-x^2} \left[-\frac{1}{2} D(-x^2)\right] dx = -\frac{1}{2} \int 10^u D_x u dx,$$

where $u = -x^2$. Hence,

$$\begin{aligned} \int x 10^{-x^2} dx &= -\frac{1}{2} 10^u \log_{10} e + C \\ &= -\frac{1}{2} 10^{-x^2} \log_{10} e + C. \end{aligned}$$

EXERCISES

Verify the following primitive functions.

1. $D^{-1}e^{-2x} = -\frac{1}{2}e^{-2x} + C$.
2. $D^{-1}(e^{-x} + e^{2x}) = -e^{-x} + \frac{1}{2}e^{2x} + C$.
3. $D^{-1}xe^{x^2} = \frac{1}{2}e^{x^2} + C$.
4. $D^{-1}xe^{-x^2} = -\frac{1}{2}e^{-x^2} + C$.
5. $D^{-1}(x+1)e^{x^2+2x} = \frac{1}{2}e^{x^2+2x} + C$.
6. $D^{-1}(\sin x e^{\cos x}) = -e^{\cos x} + C$.
7. $D^{-1}10^{x/4} = 4 \cdot 10^{x/4} \cdot \log_{10} e + C$.
8. $\int (x-2)(10^x/10^{4x}) dx = \frac{1}{2}(10^x/10^{4x}) \log_{10} e + C$.
9. $\int e^{\tan x} \sec^2 x dx = e^{\tan x} + C$.
10. $\int 3\sqrt{e^x+1} e^x dx = 2(e^x+1)^{3/2} + C$.
11. $\int e^x/(e^x+1)^2 dx = -1/(e^x+1) + C$.
12. $\int (e^x+2)^2/e^x dx = e^x + 4x - 4e^{-x} + C$.
13. $\int [e^x \log(1+e^x)]/(1+e^x) dx = 1/2[\log(1+e^x)]^2 + C$.
14. $\int e^x \cos e^x dx = \sin e^x + C$.
15. $\int (1-e^x)/(1+e^x) dx = x - 2 \log(1+e^x) + C$.

[Hint: Write $(1-e^x)/(1+e^x) = 1 - 2e^x/(1+e^x)$.]

13. The System of Functions $\int_1^x t^{n-1} dt$. Let us consider the system of equations, $y = F_n(x)$, $x > 0$, in which $F_n(x)$ is defined as follows.

$$\begin{aligned} F_n(x) &= \int_1^x t^{n-1} dt = \frac{x^n - 1}{n} \quad \text{if } n \neq 0, \\ &= \log x \quad \text{if } n = 0. \end{aligned}$$

If we plot the graphs of these equations we find that $y = \log x$ assumes a natural place in the group (Fig. 5). This suggests that the logarithmic curve is a limiting curve approached by the curves having the equations $y = (x^n - 1)/n$, as n tends to 0; and that the logarithmic function is a limiting function in this system of functions, in which all others are expressed in terms of powers of x . This is indeed true. We shall not

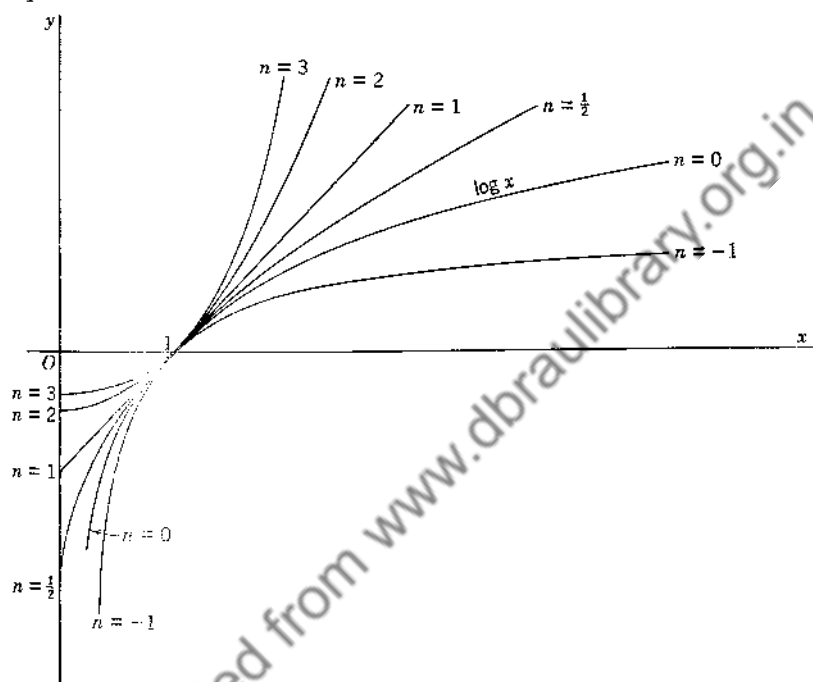


FIG. 5

prove this fact here, but instead we shall make some interesting comparisons among the functions of this system.

It is easy to see from Fig. 5 that, although all the curves pass through the point $(1, 0)$, for all other values of x the curves are arranged one above the other with increasing values of n . That is, except for the point $x = 1$, $F_p(x) > F_q(x)$ if $p > q$. The graph of $y = F_0(x)$, i.e., of $y = \log x$, divides the rest of the system into two classes: $y = F_\alpha(x) = (x^\alpha - 1)/\alpha$, where $\alpha > 0$, and $y = F_\beta(x) = (x^\beta - 1)/\beta$, where $\beta < 0$. Each of the curves of the first class rises infinitely high as $x \rightarrow \infty$, whereas each of the curves of the second class is bounded above by a horizontal asymptote, viz., $y = 1/\beta$. Also, each of the curves of the first class takes a finite value for $x = 0$, whereas the ordinates of each curve of the second class

tends to $-\infty$ as x tends to 0. Since $\log x$ tends to ∞ as x increases beyond all bounds, and tends to $-\infty$ as x tends to 0, it is natural to place the curve $y = \log x$ in the upper class for $x > 1$, as the first of the system for which the ordinate increases beyond all bounds as x does, and in the lower class for $x < 1$, as the first of this class reading downward for which the ordinate tends to $-\infty$ as x tends to 0. In the notation introduced above, F_α denotes those members of the system that are finite at $x = 0$ and become infinite with x , whereas F_β denotes those that become negatively infinite as x tends to 0 and remain finite as x increases. F_0 actually belongs to neither of these classes since it is negatively infinite as x tends to 0 and positively infinite as x tends to ∞ .

A revealing comparison of two functions both of which tend to ∞ , or both to 0, is obtained from the behavior of their ratio. We saw, for example, that although $\sin x$ tends to 0 with x , the ratio $(\sin x)/x$ tends to 1; and that $1 - \cos x$ and x^2 tend to 0 together, whereas the ratio $(1 - \cos x)/x^2$ tends to $\frac{1}{2}$. $\sin x$ is said to *vanish to the same order* as x , and $1 - \cos x$ to the same order as x^2 . In general, two functions, both of which tend to 0, vanish to the same order, or tend to 0 to the same order, if their ratio tends to some fixed number different from 0; $f(x)$ vanishes to a higher order than $g(x)$ if the ratio $f(x)/g(x)$ tends to 0, and to a lower order than $g(x)$ if $g(x)/f(x)$ tends to 0 (i.e., $f(x)/g(x) \rightarrow \infty$). Similarly, two functions, both of which tend to ∞ , do so to the same order if their ratio tends to some finite quantity different from 0; but $f(x)$ tends to ∞ to a higher order than $g(x)$ if the ratio $f(x)/g(x)$ tends to ∞ , and to a lower order than $g(x)$ if $g(x)/f(x)$ tends to ∞ (i.e., $f(x)/g(x) \rightarrow 0$). It is easy to see that x^2 tends to ∞ to a higher order than x since the ratio $x^2/x = x \rightarrow \infty$. But $3x$ and $2x$ tend to ∞ to the same order, since $3x/2x \rightarrow \frac{3}{2}$.

Among the functions F_α , it is not hard to see that $F_{\alpha_2}(x)$ tends to ∞ to a higher order than $F_{\alpha_1}(x)$ as $x \rightarrow \infty$, if $\alpha_2 > \alpha_1$. To show this, we write

$$\frac{F_{\alpha_2}}{F_{\alpha_1}} = \frac{(x^{\alpha_2} - 1)/\alpha_2}{(x^{\alpha_1} - 1)/\alpha_1} = \frac{\alpha_1(x^{\alpha_2} - 1)}{\alpha_2(x^{\alpha_1} - 1)}.$$

Dividing numerator and denominator by x^{α_1} , we have

$$\frac{F_{\alpha_2}}{F_{\alpha_1}} = \frac{\alpha_1(x^{\alpha_2 - \alpha_1} - x^{-\alpha_1})}{\alpha_2(1 - x^{-\alpha_1})}.$$

Since $\alpha_2 > \alpha_1 > 0$, we have $\alpha_2 - \alpha_1 > 0$ and $-\alpha_1 < 0$. Hence, $x^{\alpha_2 - \alpha_1} \rightarrow \infty$ and $x^{-\alpha_1} \rightarrow 0$ as $x \rightarrow \infty$. Therefore, $F_{\alpha_2}/F_{\alpha_1}$ must tend to ∞ , since the numerator tends to ∞ while the denominator remains finite.

We can now show that F_α tends to ∞ to a higher order than $\log x$, as

$x \rightarrow \infty$, for every $\alpha > 0$. We note first that if α_1 is any number greater than 0, and in particular if $\alpha > \alpha_1 > 0$, then $F_{\alpha_1} > \log x$ for all values of x greater than 1. This is easily seen from the graph and follows from the definitions of these functions by the integral. Therefore,

$$F_{\alpha}/\log x > F_{\alpha}/F_{\alpha_1}.$$

Since we have just shown that $F_{\alpha}/F_{\alpha_1} \rightarrow \infty$ as $x \rightarrow \infty$, it follows that $F_{\alpha}/(\log x) \rightarrow \infty$ as $x \rightarrow \infty$. That is,

$$\frac{(x^{\alpha} - 1)/\alpha}{\log x} \rightarrow \infty \quad \text{as } x \rightarrow \infty. \quad (13.1)$$

This is true for every $\alpha > 0$, since a value of α_1 , such that $\alpha > \alpha_1 > 0$, can always be found.

14. Comparison of $\log x$, x^n , and a^x . As a corollary to (13.1), it follows that the ratio $x^{\alpha}/(\log x) \rightarrow \infty$ as $x \rightarrow \infty$, if $\alpha > 0$. This can be seen by writing

$$\frac{x^{\alpha}}{\log x} = \frac{x^{\alpha}}{(x^{\alpha} - 1)/\alpha} \cdot \frac{(x^{\alpha} - 1)/\alpha}{\log x},$$

and observing that

$$\lim_{x \rightarrow \infty} \frac{x^{\alpha}}{(x^{\alpha} - 1)/\alpha} = \lim_{x \rightarrow \infty} \left(\alpha \cdot \frac{1}{1 - 1/x^{\alpha}} \right) = \alpha.$$

Hence, since the fraction $\frac{(x^{\alpha} - 1)/\alpha}{\log x} \rightarrow \infty$, by (13.1), so must $x^{\alpha}/(\log x)$.

Thus, any positive power of x tends to ∞ to a higher order than $\log x$. That is,

$$\frac{x^n}{\log x} \rightarrow \infty \quad \text{as } x \rightarrow \infty, \text{ if } n > 0. \quad (14.1)$$

By inverting this ratio, we conclude that

$$\frac{\log x}{x^n} \rightarrow 0 \quad \text{as } x \rightarrow \infty, \text{ if } n > 0. \quad (14.2)$$

It is natural to ask whether a similar conclusion can be drawn regarding the orders to which $\log x$ and F_{β} tend to $-\infty$, as x tends to 0. The situation as $x \rightarrow 0$ is brought within the scope of the above discussion by a transformation of the expression for $F_n(x)$ made by putting $x = 1/\xi$ ($\xi =$ Greek Xi), so that $\xi \rightarrow \infty$ as $x \rightarrow 0$. We then have

$$\begin{aligned} F_n(x) &= (\xi^{-n} - 1)/n, \quad \text{if } n \neq 0 \\ &= \log(1/\xi) = -\log \xi, \quad \text{if } n = 0. \end{aligned}$$

Now,

$$\frac{F_\beta}{\log x} = \frac{(x^\beta - 1)/\beta}{\log x} = \frac{(\xi^{-\beta} - 1)/\beta}{-\log \xi},$$

wherein $\beta < 0$, and $-\beta > 0$. As $x \rightarrow 0$ and $\xi \rightarrow \infty$, this ratio has the form of (13.1). It therefore tends to ∞ as $\xi \rightarrow \infty$, i.e., as $x \rightarrow 0$.

As a corollary we have

$$x^\beta/(\log x) \rightarrow -\infty \text{ as } x \rightarrow 0, \text{ if } \beta < 0. \quad (14.3)$$

The negative sign results from the occurrence of the negative number β in $(x^\beta - 1)/\beta$. Thus,

$$\frac{x^\beta}{\log x} = \frac{x^\beta}{(x^\beta - 1)/\beta} \cdot \frac{(x^\beta - 1)/\beta}{\log x}.$$

Now, the first factor has the negative limit β , and the second tends to ∞ . Their product must then tend to $-\infty$. The result in (14.3) can be written

$$\frac{x^{-n}}{\log x} \rightarrow -\infty \text{ as } x \rightarrow 0, \text{ if } n > 0, \quad (14.4)$$

whence

$$x^n \log x \rightarrow 0 \text{ as } x \rightarrow 0, \text{ if } n > 0. \quad (14.5)$$

The comparison of the orders of exponential functions and powers of x is brought out in Exercises 6–10, below.

EXERCISES

1. Prove the corollary stated in (14.4) by transforming the fractional form that occurs in (14.1).
2. Find $\lim_{x \rightarrow \infty} x/(\log x^{100})$.
3. Find $\lim_{x \rightarrow 0} \sqrt[n]{x} \log x$.
4. Find $\lim_{x \rightarrow 0} \sin x \cdot \log x$. [Hint: Use $\lim_{x \rightarrow 0} (\sin x)/x = 1$.]
5. Find $\lim_{x \rightarrow 0} \sqrt{1 - \cos x} \cdot \log x$. [Hint: Use $\lim_{x \rightarrow 0} (1 - \cos x)/x^2 = \frac{1}{2}$.]
6. Prove that $e^x/x \rightarrow \infty$ as $x \rightarrow \infty$. [Hint: Let $x = \log \xi$.]
7. Prove that $e^{\alpha x}/x^\beta \rightarrow \infty$ as $x \rightarrow \infty$, if α and β are any positive numbers.
8. Prove that $xe^{-x} \rightarrow 0$ as $x \rightarrow \infty$.
9. Prove that $xe^a \rightarrow 0$ as $x \rightarrow \infty$ if $0 < a < 1$.
10. Prove that $nx^{n-1} \rightarrow 0$ as $n \rightarrow \infty$ if $0 < x < 1$.
11. Find $\lim_{x \rightarrow \infty} x^n e^{-x}$.
12. Find $\lim_{x \rightarrow \infty} xe^{-x} \log x$.

13. Show that $\lim_{x \rightarrow \infty} a^{1/x} = 1$, and hence deduce $\lim_{n \rightarrow \infty} \sqrt[n]{a}$, where a is a positive number.
14. Show that $\lim_{x \rightarrow 0} x^x = 1$. [Hint: Write $x^x = (e^{\log x})^x = e^{x \log x}$.]
15. Show that $\lim_{x \rightarrow \infty} x^{1/x} = 1$. From this result deduce the value of $\lim_{n \rightarrow \infty} \sqrt[n]{n}$.

15. The Differential Equation $D_x y = ay$. We saw that the exponential function e^x has its derived function equal to itself. This function thus satisfies the differential equation $D_x y = y$. The same is true of the function ce^x , since $D_x ce^x = ce^x$.

The differential equation $D_x y = ay$ must be satisfied by a function that has its derived function proportional to itself, with the proportionality factor equal to a . Clearly e^{ax} is such a function, since $D_x e^{ax} = ae^{ax}$. A more general solution is ce^{ax} , where c is an arbitrary constant. A theorem in the theory of differential equations shows that every solution of the differential equation

$$D_x y = ay \quad (15.1)$$

has the form

$$y = ce^{ax}.$$

The statement that Eq. (15.1) is satisfied by a quantity that changes at a rate that, at any instant, is proportional to the quantity itself at that instant, is a trivial remark, since, this statement only translates the equation into words. It may serve, however, to direct attention toward some practical applications. One of the first to come to mind is that of an amount of money at compound interest. The interest for any period is a percentage of the amount at the beginning of that period. This does not fit a differential equation like (15.1) exactly, since the amount is not a continuous function of the time. It increases by a distinct jump at the end of each interest period. If the interest were added in continuously as it accrued, at the rate of 3 per cent per year, the differential equation satisfied by the growing amount would be

$$D_t A = .03A,$$

where t denotes the time elapsed in years. The differential equation is satisfied by $A = ce^{.03t}$. In this result, c stands for the principal at the start, since when $t = 0$, $A = c$. The presence of an arbitrary constant c in the solution reflects the obvious fact that the differential equation must apply equally well for any initial sum. If the original principal were \$100, the solution would be $A = 100e^{.03t}$.

Let us consider another example.

EXAMPLE. Radium decomposes continuously at a rate per century that is always equal to 3.8 per cent of the amount Q that remains. Write

the differential equation for Q , and express the amount Q remaining at the end of t centuries, if the present quantity is 50 mg. How much will be left at the end of 1000 yr? When will exactly one-half of the radium be left?

SOLUTION. We have $D_t Q = -.038 Q$ as the differential equation for Q . The rate is negative since Q decreases. The general solution of this differential equation is

$$Q = ce^{-.038t}.$$

Since the present is represented by $t = 0$, we have

$$Q = 50, \quad \text{when} \quad t = 0.$$

Hence, $50 = ce^0$, whence $c = 50$, and for any value of t , $Q = 50e^{-.038t}$.

In order to find Q at the end of 1000 yr, we set $t = 10$, since 1000 yr is equal to 10 centuries. At this time,

$$Q = 50e^{-.038 \cdot 10} = 50e^{-.38}.$$

From the table of values of e^{-x} we find that $e^{-x} = .684$ when $x = .38$. Hence, $Q = 50 \cdot .684 = 34.2$. This is the number of mg left at the end of 1000 yr.

A problem of practical importance is that of finding the time when one-half the initial amount of a radioactive substance is left—the *half-life*. For this purpose we set

$$Q = \frac{1}{2}(50) = 25, \quad \text{or} \quad 50e^{-.038t} = 25,$$

and solve for t . We have $e^{-.038t} = .5$, from which

$$-.038t = \log_e .5 = -.6931.$$

Hence,

$$t = .6931/.038 = 18.24.$$

The half-life is therefore 18.24 centuries, or 1824 yr.

EXERCISES

Find y as a function of x satisfying the following conditions.

1. $D_x y = 2y$; $y = 3$ when $x = 0$.
2. $D_x y = -y$; $y = 10$ when $x = 0$.
3. $D_x y + 4y = 0$; $y = 1/e$ when $x = \frac{1}{4}$.
4. $D_x y - .3y = 0$; $y = e$ when $x = 10$.
5. $10D_x y = y$; $y = 2.83$ when $x = 0$.
6. Find the equation of the curve having its slope always equal to $\frac{1}{x}$ the ordinate, and crossing the y -axis at $(0, 3)$.
7. Find the equation of the curve having its slope always equal to the negative of the ordinate and passing through $(1, 2/e)$.

8. Find the amount of \$10,000 after 20 yr if it is compounded continuously at the rate of 3 per cent per year.

9. An electric current of i amperes died out in such a way that t seconds after the voltage was cut off

$$D_t i = -50i.$$

If $i = 20$ when $t = 0$, express i as a function of t .

10. If p is the pressure of the atmosphere, measured by the number of inches of mercury in a barometer, and if h is the height in miles above sea level, then

$$D_h p = -.2p.$$

(a) Express p as a function of h if the pressure at sea level is 29.9 in. (b) Find the pressure 2 miles above sea level. (c) At what height is the pressure 15 in.?

11. The number Q of bacteria in a culture grows continuously at a rate per hour that is 4.3 per cent of the number present. Write the differential equation for Q , and express Q as a function of the number of hours, t , if there are Q_0 at the start. How many bacteria are present at the end of 4 hr?

12. The temperature of the sun, T degrees, decreases continuously at a rate per million years which is 20 per cent of T . Write the differential equation for T , and express T as a function of the number of million-year periods elapsed if T is now 12,000. What will the temperature be in 10 million years? When will the temperature be half what it is today?

16. Definite Integrals. Some new definite integrals are now at our command as a result of the formulas for primitive functions derived in this chapter. These will be illustrated in examples.

EXAMPLE 1. Find the area bounded by the rectangular hyperbola $xy = 1$, the x -axis, and the ordinates $x = 1$ and $x = a$, where $a > 1$.

SOLUTION. We have

$$\text{area} = \int_1^a \frac{1}{x} dx = \log a - \log 1 = \log a.$$

More generally, if the ordinates are $x = a$ and $x = b$, where $0 < a < b$,

$$\text{area} = \int_a^b (1/x) dx = \log b - \log a = \log b/a. \quad (\text{See Exercise 21.})$$

EXAMPLE 2. Find the volume generated by revolving about the x -axis the area bounded by the exponential curve $y = e^{-x}$, the coordinate axes and the ordinate $x = 2$.

SOLUTION. We have

$$\begin{aligned} V &= \int_0^2 \pi(e^{-x})^2 dx = \pi \int_0^2 e^{-2x} dx \\ &= \pi \left[-e^{-2x}/2 \right]_0^2 = \pi(\frac{1}{2} - e^{-4}/2) \\ &= \pi(.5 - .00916) = .49084\pi. \end{aligned}$$

EXAMPLE 3. Find the area bounded by the catenary $y = \frac{1}{2}(e^{x/2} + e^{-x/2})$, the x -axis, and the ordinates $x = -1$ and $x = 1$ (Fig. 6).

SOLUTION. The catenary is the curve naturally assumed by a perfectly flexible heavy cable hanging freely between two supports (see also Chapter 18, Art. 6). Because of the symmetry of the catenary about the y -axis, the required area is

$$\begin{aligned} A &= 2 \int_0^1 \frac{1}{2}(e^{x/2} + e^{-x/2}) dx \\ &= 2 \left[e^{x/2} - e^{-x/2} \right]_0^1 = 2(e^{\frac{1}{2}} - e^{-\frac{1}{2}}) \\ &= 2(1.649 - .607) = 2.084. \end{aligned}$$

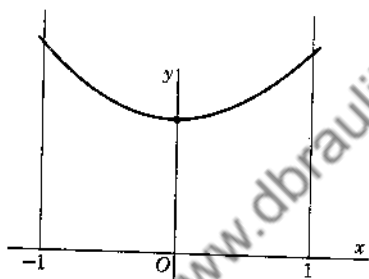


FIG. 6

EXERCISES

Calculate the following definite integrals.

1. $\int_1^{10} \frac{1}{x} dx.$
2. $\int_{.1}^1 \frac{1}{x} dx.$
3. $\int_{-10}^0 e^x dx.$
4. $\int_{-1}^0 10^x dx.$
5. $\int_1^e (x - 1/x) dx.$
6. $\int_0^1 \frac{x}{x^2 + 1} dx.$
7. $\int_0^{\pi/3} \tan x dx.$
8. $\int_{-\pi/4}^{\pi/4} \sec x dx.$
9. $\int_1^e \frac{\log x}{x} dx.$
10. $\int_{\pi/4}^{\pi/2} \cot x \log \sin x dx.$
11. $\int_0^2 x e^{-\frac{1}{2}x^2} dx.$
12. $\int_2^{10} \frac{1}{x \log_{10} x} dx.$

Find the areas bounded by the following curves, the x -axis, and the ordinates specified.

13. $y = \tan x$; $x = 0$, $x = \pi/4$.

14. $y = \tan x$; $x = 0$, $x = 22\pi/45$.

15. $y = \frac{1}{x} - 1$; $x = \frac{1}{2}$, $x = 1$.

16. $y = e^{-.1x}$; $x = 0$, $x = 10$.

17. $y = \frac{1}{x-1}$; $x = 0$, $x = 9$.

18. $y = \frac{x}{x^2 - 1}$; $x = \sqrt{2}$, $x = 5$.

19. $y = 10^x$; $x = -2$, $x = 0$.

20. $y = \sec x$; $x = -\pi/3$, $x = \pi/3$.

21. Show that the area bounded by the rectangular hyperbola $xy = k$, the x -axis, and the two ordinates $x = a$ and $x = ba$ is the same for all values of a if b is fixed, and assuming that a and b are both positive.

22. Find the area bounded by the curve $xy = 1$, the y -axis, and the lines $y = 1$, and $y = 2$.

23. Find the area bounded by the curve $y = \log x$, the coordinate axes, and the line $y = 1$.

24. Find the area bounded by the curve $xy = 3$ and the straight line $x + y = 4$.

25. Find the volume generated by revolving about the y -axis the area bounded by the curve $y = \log x$, the coordinate axes, and the line $y = -2$.

26. Find the volume generated by revolving about the x -axis the area bounded by the curve $xy^2 = 1$, the x -axis, and the lines $x = \frac{1}{2}$ and $x = 2$.

27. Evaluate $\int_0^1 e^{x^2} dx$ by direct application of the definition of the definite integral as

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i) \Delta x = \lim_{n \rightarrow \infty} \sum_{i=1}^n e^{x_i^2} \Delta x.$$

[Hint: $\sum_{i=1}^n e^{x_i^2} \Delta x = \sum_{i=1}^n e^{ih} h$, where $h = 1/n$

$$= (e^h + e^{2h} + e^{3h} + \cdots + e^{n^2 h})h.]$$

Differentials. The Mean Value Theorem

1. The Problem of Approximating a Functional Value. The most obvious use of a functional relation is the calculation of values of the function corresponding to particular values of the independent variable. This is a comparatively simple matter with many of the functions that we



FIG. 1

have studied. It offers great difficulties in the case of many more advanced functions, particularly functions beyond the scope of this book. Some functions in an intermediate range, more complicated than most of the simple algebraic functions we have studied thus far but much more easily evaluated than many functions of advanced analysis, are easily evaluated for some values of the independent variable but offer more difficulty for others. For example, the function $f(x) = \sqrt[5]{x}$ has the value 3 if $x = 243$ (this is easily found by experiment). A considerable amount of labor, experimental or otherwise, is required to find its value for $x = 250$, or for 380.

In practical work we are usually satisfied with approximations to values that are difficult or impossible to evaluate exactly, provided we can be sure that the error lies within certain tolerable limits. The problem of approximation of functions therefore has great practical importance.

Let us consider the problem of approximating $f(x) = \sqrt[5]{x}$ for $x = 250$, i.e., $\sqrt[5]{250}$, from a graphical point of view. Figure 1 shows the graph of $y = \sqrt[5]{x}$ for $x > 0$. Of course, if we have an accurate graph of this function, the value of $\sqrt[5]{x}$, for any value of x , can be read from the graph. But we could not expect to get accurate results this way. Moreover, the graph itself could not be drawn without the knowledge of such values as

we are trying to get. Here, the graph is used merely as a means of guiding us to a numerical method. For that purpose it need not be drawn with great exactness.

Because of the large values of x in which we are interested, it will be well to consider a small part of the graph in Fig. 1 near $x = 250$, in a magnified scale. This is shown in Fig. 2. The length of line MP represents $\sqrt[5]{243} = 3$. The length of line NQ represents $\sqrt[5]{250}$, and it is this that we are to find approximately. Since $NQ = NS + SQ = MP + SQ$, and MP is known, if we could find SQ our problem would be solved.

In order to approximate SQ we might draw the tangent at P and use SR , which is the rise of the tangent in the horizontal distance PS , as an

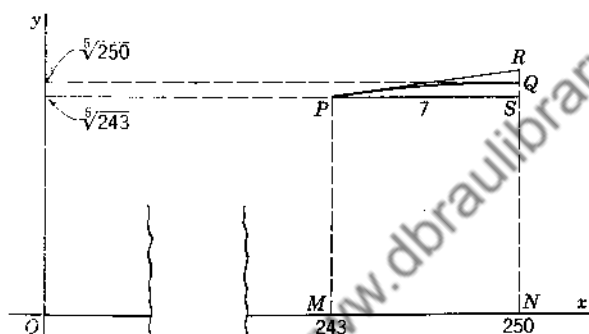


FIG. 2

approximation to the rise of the curve. It is natural to try to approximate the points of the curve by means of the simplest of all graphs, namely a straight line; and when one considers all the points of the curve in the immediate neighborhood of P , the tangent appears to approximate these points more closely than any other straight line.

The slope of the tangent to a curve $y = f(x)$, at $P(x_0, y_0)$, is simply $f'(x_0)$. In the example in question

$$f(x) = \sqrt[5]{x} \quad \text{and} \quad f'(x) = 1/[5(\sqrt[5]{x})^4].$$

Also,

$$x_0 = 243, \text{ so that } f'(x_0) = 1/[5(\sqrt[5]{243})^4] = .00247.$$

But the slope is also equal to $SR/PS = SR/7$. Hence,

$$SR/7 = .00247, \quad \text{and} \quad SR = 7(.00247) = .017.$$

We can then write

$$\sqrt[5]{250} = \sqrt[5]{243} + 7 = MP + SQ \approx MP + SR = 3 + .017 = 3.017.$$

The symbol $\approx$ is read "is approximately equal to." The value of $\sqrt[5]{250}$ found by means of 7-place logarithms is 3.01701.

It is apparent from the graph that we cannot use this method to approximate $\sqrt[5]{380}$ and expect to get a satisfactory result, since at $x = 380$ the tangent at P will be at a considerable distance from the curve. We would need to know the ordinate and slope of the tangent at a point much nearer 380 than is 243. These quantities could not be conveniently found.

This method of approximating a function at a point near a value for which the function is known will be systematized and facilitated by means of some new notations introduced in Art. 2.

2. Differences and Differentials. Our study of the rate problem led us to consider two values of the independent variable differing by h , and two corresponding values of the function, whose difference we may denote by k . These quantities, h and k , are often called *differences*, and the quotient k/h which was used in finding the derivative, is called the *difference quotient*. The derivative $= \lim_{h \rightarrow 0} k/h$, and is not a difference quotient; it is not found by a mere division, but rather by applying a limit process to a function that is found by division.

It is often convenient to use symbols for differences that indicate the variables to which they apply. Thus, if $y = f(x)$, in place of h we use Δx (delta- x). Similarly, in place of k , we use Δy . One must note carefully that Δx and Δy are single quantities and not products. With these symbols,

Δx = difference of x ,

Δy = difference of y , or $\Delta f(x)$ = difference of $f(x)$,

$\frac{\Delta y}{\Delta x} = \frac{\Delta f(x)}{\Delta x}$ = difference quotient,

and

$$Dy = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}.$$

The difference quotient is also written

$$\frac{y_2 - y_1}{x_2 - x_1}, \frac{f(x_2) - f(x_1)}{x_2 - x_1}, \text{ and } \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x}.$$

The basic steps leading to the derivative are (1) the determination of a difference Δy corresponding to an arbitrary difference Δx , (2) the formation of the quotient $\Delta y/\Delta x$, and (3) the calculation of $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$. We now use the final result, found in any way, e.g., by formula, to get an approximation to a new difference of the function which corresponds to a new arbitrary difference of the independent variable. This was the program

followed in Art. 1. The exact difference was difficult to compute, whereas the approximation is easily obtained.

If $y = f(x)$, the new arbitrary difference of x is denoted by dx , called the *differential* of x . The approximate difference of the function is denoted by dy (or $df(x)$) and is called the differential of y (or of $f(x)$). If, as in Art. 1, we know the value of $f(x)$ at a point $x = x_1$, we find the approximate difference of $f(x)$, dy , corresponding to an arbitrary difference of x , dx , by means of the formula,

$$dy = f'(x_1) dx. \quad (2.1)$$

In this relation dy is a linear function of dx , and represents the change in height of the tangent due to a change of dx in x . In general, when the value of x is not specified, dy is a function of two independent variables, x and dx ,

$$dy = f'(x) dx. \quad (2.2)$$

In the example in Art. 1, we approximated the value of SQ , which is a particular value of the difference of y , Δy , corresponding to the difference of x , $dx = 7$. Since dx and Δx are both differences of x there is no harm in identifying them in the sense that any particular value of Δx can be taken as a value of dx . In fact, the value of dx that yields a definite dy is equal to the value of Δx that corresponds to the value of Δy that is approximated by dy . The fact that we think of Δx as tending to 0 in the calculation of the derivative, however, should not be carried over to dx , since dx is an arbitrary variable that may or may not be small.

3. Approximations by Means of Differentials. The approximation in Art. 1 was found by means of a differential, although this term was not used there. We approximated $\sqrt[5]{250}$ by the following method. We set

$$y = \sqrt[5]{x}, \quad y + \Delta y = \sqrt[5]{x + \Delta x}.$$

Putting

$$x = 243, \quad dx = \Delta x = 7,$$

we have

$$y = 3, \quad y + \Delta y = \sqrt[5]{250}.$$

Since we seek the value of $y + \Delta y$ and we have $y = 3$, it is only necessary to determine Δy . This we approximate by dy , writing $dy \approx \Delta y$. Now, since

$$Dy = \frac{1}{5(\sqrt[5]{x})^4}, \quad dy = \frac{1}{5(\sqrt[5]{x})^4} dx.$$

Inserting $x = 243$, $dx = 7$ in this expression for dy ,

$$dy = \frac{1}{5(\sqrt[5]{243})^4} \cdot 7 = \frac{7}{5 \cdot 81} = \frac{7}{405} = .017.$$

Hence,

$$y + \Delta y \approx y + dy = 3 + .017 = 3.017.$$

Let us consider some additional examples.

EXAMPLE 1. Use a differential to find the approximate decrease in the function $1/x^2$ as x changes from 1 to 1.01.

SOLUTION. We shall find dy as an approximation to Δy in the function $y = 1/x^2$, with $x = 1$, $y = 1/1^2 = 1$, $dx = .01$. Differentiating,

$$dy = -\frac{2}{x^3} dx.$$

Inserting $x = 1$, $dx = .01$,

$$dy = -\frac{2}{1^3} (.01) = -.02.$$

Hence,

$$\Delta y \approx -.02.$$

Thus, the decrease is approximately .02 units and the value of the function is then $y + \Delta y \approx y + dy = 1 - .02 = .98$.

EXAMPLE 2. The area of a square is found by measuring the side x and using the formula $A = x^2$. The side is found to be 12 in. If this measurement is in error by .1 in., find the approximate error in the computed area, and the approximate relative error.

SOLUTION. We have $A = x^2$, and $dA = 2x dx$. Letting

$$x = 12, \quad dx = \pm .1,$$

$$dA = 2 \cdot 12 \cdot (\pm .1) = \pm 2.4.$$

We use the double sign $\pm$ because the error may be positive or negative, that is, the measurement may be too much or too little. The amount of the error is therefore approximately 2.4 sq. in.

Relative error is defined to be the ratio of the error to the true value. It can usually be closely approximated by the ratio of the error to the measured or computed value. For our approximation to the relative error we shall take dA/A . The computed value of A is $A = 12^2 = 144$. Therefore, approximate relative error $= 2.4/144 = .0166 \dots$. The relative error can be found more expeditiously from $A = x^2$ by first changing this equation into a logarithmic equation

$$\log A = 2 \log x.$$

Now, differentiating,

$$D_x A/A = 2/x, \quad \text{and} \quad (D_x A dx)/A = 2 dx/x.$$

Since $D_x A dx = dA$,

$$dA/A = 2 dx/x. \quad (3.1)$$

In view of the discussion in Art. 6 below, $\log A$ can be differentiated to produce dA/A just as though A were the independent variable. Inserting $x = 12$, $dx = .1$ in (3.1), we have

$$dA/A = 2(.1)/12 = .0166 \dots$$

EXAMPLE 3. Approximate the value of $\sin 31^\circ$ by means of differentials, given $\sin 30^\circ = .5$ and $\cos 30^\circ = .86603$.

SOLUTION. Let

$$y = \sin x.$$

Then,

$$dy = \cos x \, dx.$$

Since $\sin 30^\circ = \sin (\pi/6) = .5$, $\cos 30^\circ = \cos (\pi/6) = .86603$, and $1^\circ = .01745$ rad,* approximately, we put

$$x = \pi/6, \quad y = .5, \quad dx = .01745.$$

We shall approximate $y + \Delta y$ by $y + dy$. We therefore calculate $dy = (.86603)(.01745) = .01511$, approximately. Using this value of dy as an approximation to Δy , we have

$$y + \Delta y \approx .5 + .01511 = .51511.$$

The value of $\sin 31^\circ$ given in a 5-place table is .51504.

EXERCISES

1. Use differentials to approximate the change in each of the following functions that corresponds to the indicated change of the independent variable.

(a) x^3 as x changes from 4 to 4.02.

(b) $x^2 + x - 10$ as x changes from 10 to 9.98.

(c) $1/x$ as x changes from 1 to 1.03.

(d) $\sqrt{x^2 + 25}$ as x changes from 12 to 11.5.

(e) $\log_{10} x$ as x changes from 100 to 102, given $\log_{10} e = .4343$.

(f) $\tan x$ as x changes from 45° to 44° .

2. Approximate the following radicals by means of differentials.

(a) $\sqrt{99}$.

(b) $\sqrt{147}$.

(c) $\sqrt[3]{66}$.

(d) $\sqrt[4]{17}$.

(e) $\sqrt{162}$.

(f) $\sqrt[3]{8.05}$.

3. Approximate the following quantities by means of differentials.

(a) 17^2 .

(b) $1/\sqrt{47}$.

(c) $\log_{10} 990$.

(d) $\cos 28^\circ$.

(e) $\sin 59^\circ$.

(f) $e^{1.2}$.

* It is necessary to use radian measure because $D \sin x$ is equal to $\cos x$ only if $x =$ the number of radians in the angle.

4. A circular plate expands as it is heated so that its radius increases from 6 in. to 6.08 in. Find the approximate increase in area.
5. Use differentials to find the approximate volume of a spherical shell with its inner radius 6 in. and its outer radius 6.1 in.
6. The period, T seconds, of a simple pendulum (time for a complete oscillation) is given by $T = 2\pi\sqrt{l/g}$, in which l is the length of the pendulum in feet and g is the acceleration due to gravity in feet per second per second. Using $g = 32$, find the approximate change in T due to increasing l from 2 ft to 2.1 ft.
7. A sphere of radius 5 in. shrinks to radius 4.8 in. Find the approximate decrease in volume and surface area.
8. Under ideal conditions, according to Boyle's law, the pressure P pounds and volume V cubic inches of a perfect gas satisfy the equation $PV = K$ — a constant. If $K = 60$, and the pressure is found by measurement to be 1.5 lb with a maximum error of .05 lb, approximate the maximum error and relative error in the calculated value of V .

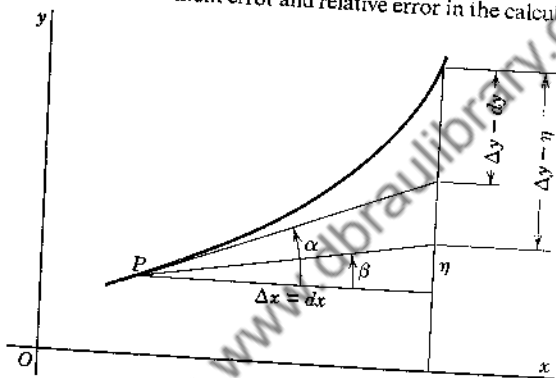


FIG. 3

4. Differentiable Functions. It is intuitively apparent that the tangent to a curve, $y = f(x)$, at a point $P(x_0, y_0)$, approximates the curve more closely in a neighborhood of P than does any other straight line. We can prove this by comparing the error with that of any other linear approximation. The error of the approximation by the tangent is $\Delta y - dy$. Obviously, any other approximating line must pass through P . Let such a line have the inclination $\beta \neq \alpha$, where $\tan \alpha = f'(x_0)$. If we denote its ordinate at $x_0 + \Delta x$ by $y_0 + \eta$, the error is $\Delta y - \eta$ (Fig. 3). Now, both $\Delta y - dy$ and $\Delta y - \eta$ tend to 0 with Δx , but we can show that $\Delta y - dy$ vanishes to a higher order than $\Delta y - \eta$ (see Art. 14, Chapter 11). Hence, $\Delta y - dy$ must be less than $\Delta y - \eta$ for sufficiently small values of Δx . We have

$$dy = f'(x_0)\Delta x = \tan \alpha \cdot \Delta x.$$

Since $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(x_0)$, from the definition of a limit, we have

$$\frac{\Delta y}{\Delta x} = f'(x_0) + \epsilon$$

and

$$\Delta y = f'(x_0)\Delta x + \varepsilon\Delta x = \tan \alpha \cdot \Delta x + \varepsilon\Delta x,$$

in which

$$\varepsilon > 0 \text{ as } \Delta x \rightarrow 0, \text{ i.e., as } x \rightarrow x_0.$$

Now, as we see from Fig. 3,

$$\eta = \tan \beta \cdot \Delta x.$$

Hence,

$$\frac{\Delta y - dy}{\Delta y - \eta} = \frac{\varepsilon\Delta x}{(\tan \alpha - \tan \beta)\Delta x + \varepsilon\Delta x} = \frac{\varepsilon}{\tan \alpha - \tan \beta + \varepsilon}.$$

Since $\tan \alpha - \tan \beta$ is a fixed number different from 0, it follows that $(\Delta y - dy)/(\Delta y - \eta) \rightarrow 0$ as $\Delta x \rightarrow 0$. We can therefore say that the error, $\Delta y - dy$, is small in comparison with $\Delta y - \eta$, when Δx is sufficiently small. Consequently, the approximation by the tangent is better than that by any other straight line.

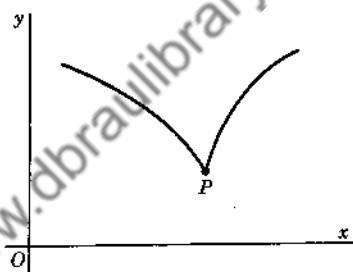


FIG. 4

A curve like that shown in Fig. 4 has no true tangent at point P . For this curve there is no best linear approximation in a neighborhood surrounding P , i.e., one for which the ratio of the error to that of any other linear approximation tends to 0 with Δx . Functions for which the differential dy gives a best approximation to Δy are said to be *differentiable*. The function illustrated in Fig. 4 is not differentiable at P . Differentiability and the existence of a tangent line are seen to be equivalent. Likewise, differentiability of a function and its possession of a derivative are equivalent. Differentiability of a function is a property that may belong to a function at some points and be denied to it at others. Fortunately, most of the functions that we shall encounter are differentiable everywhere except at some isolated points, which can be handled without great difficulty.

5. The Symbols dy/dx and d/dx . It follows, as a consequence of the definitions of the differentials dy and dx , that the quotient $dy \div dx$ is equal to the derivative $f'(x)$. That is,

$$dy/dx = D_x y = f'(x).$$

It should be noticed that in order to find dy we must first find $f'(x)$, since $dy = f'(x)dx$. We do not find the derivative by dividing the differentials: $dy \div dx$. On the contrary, we find the differential dy by multiplying the

derivative $f'(x)$ by the differential dx . Although in some of the examples of Art. 6 it will appear that the derivative is found by the division of differentials, it will be found that derivatives are involved in the calculation of the differentials, and the manipulation of the differentials is only a convenient device for obtaining the results of corresponding manipulation of the derivatives from which they are obtained.

The symbol dy/dx may be thought of as the quotient of the differentials which is equal to the derivative, or as the derivative itself. The latter usage is very common, since in many treatments of the subject of calculus this symbol is used for the derivative from the start. If one expresses the difference quotient by $\Delta y/\Delta x$, the limit of $\Delta y/\Delta x$ as $\Delta x \rightarrow 0$ is then denoted by dy/dx , passage to the limit being indicated by the change from Δ to d . When this is done, the symbol dy/dx is one inseparable symbol and not a fraction, even though $\Delta y/\Delta x$ is a fraction. This symbol for the derivative is often written in a slightly different form, $\frac{d}{dx}y$. Here d/dx denotes the operation of differentiation applied to y . This symbol can be applied to any function of x ; for example,

$$\frac{d}{dx}f(x), \quad \frac{d}{dx}(x^2 + 3x \sin x).$$

The two symbols, d/dx and D_x , are identical in meaning. We shall use $\frac{dy}{dx}$ and $\frac{d}{dx}f(x)$ frequently in the remainder of this book, since these notations are in common usage in other books, both elementary and advanced, and since some advantages result from their use that will be noted as encountered.

6. Differential Formulas and Their Use. From the definition of the differential of a function it is clear that we need no new formulas in order to find differentials of all functions whose derivatives we can find. For example, since

$$\frac{d(x^n)}{dx} = \frac{d}{dx}x^n = nx^{n-1}, \quad d(x^n) = nx^{n-1}dx.$$

There is one differential formula, however, that has a special interest, namely, the formula for the differential of a compound function. Let

$$y = f(u), \quad u = g(x).$$

Then, by the chain rule,

$$D_x y = dy/dx = f'(u)g'(x).$$

Hence,

$$dy = f'(u)g'(x) dx. \quad (6.1)$$

Since the differential of u is

$$du = g'(x) dx,$$

Eq. 6.1 can be written

$$dy = f'(u) du. \quad (6.2)$$

If u were the independent variable, and not an intermediate variable dependent upon x , Eq. 6.2 would still be the correct formula for dy . Thus, Eq. 6.2 holds whether or not u is independent. Differentials owe much of their usefulness to this fact, because of which it is unnecessary to determine in advance what the independent variable is.

The derivative of a function depends upon what variable is independent. That is, the derivative of a function with respect to one variable need not be equal to its derivative with respect to a different variable. Thus, for example, if $y = 3u$ and u is independent,

$$dy/du = 3.$$

But, if u is itself dependent on x , through the equation $u = 4x$, then

$$dy/dx = dy/du \cdot du/dx = 3 \cdot 4 = 12.$$

Since the derivative is the coefficient of the differential of the independent variable, this coefficient depends upon what variable is independent. This is not a contradiction of Eq. 6.2, which merely states that the same form can be used to obtain dy , whether u is independent or $u =$ a function of x . In the example just given,

$$y = f(u) = 3u, \quad \text{and} \quad dy = f'(u) du = 3du$$

$$y = \phi(x) = 12x, \quad \text{and} \quad dy = \phi'(x) dx = 12dx.$$

The forms of $f'(u) du$ and $\phi'(x) dx$ are the same, but $f'(u)$ and $\phi'(x)$ have different values. The fact is, of course, that $du = 4dx$, and hence both results are the same when expressed in terms of the same variable.

Since, in the expression for the differential of a function, the coefficient of the differential of the independent variable is always the derivative of the function, the derivative is often called the *differential coefficient*.

We give some examples of differentials and their use in connection with finding derivatives.

EXAMPLE 1. Find dy if $y = x^2 - 1/x^2$.

SOLUTION. We have

$$Dy = 2x + 2/x^3.$$

Hence,

$$dy = (2x + 2/x^3) dx.$$

EXAMPLE 2. Find dy if $y = \sqrt[n]{x}$.

SOLUTION. Writing $\sqrt[n]{x} = x^{\frac{1}{n}}$,

$$\begin{aligned} Dy &= \frac{1}{n} x^{\frac{1}{n}-1} \\ &= \frac{1}{n} x^{\frac{1-n}{n}} = \frac{1}{n} x^{-\frac{n-1}{n}} \\ &= \frac{1}{nx^{\frac{n-1}{n}}} = \frac{1}{n\sqrt[n]{x^{n-1}}}. \end{aligned}$$

Hence,

$$dy = \frac{1}{n\sqrt[n]{x^{n-1}}} dx.$$

EXAMPLE 3. Find dy if $y = \sqrt{x^2 - 1}$, $x = 1 - 2t^2$.

SOLUTION.

$$dy = \frac{x}{\sqrt{x^2 - 1}} dx \quad (6.3)$$

and

$$dx = -4t dt.$$

Therefore,

$$\begin{aligned} dy &= -\frac{4xt}{\sqrt{x^2 - 1}} dt, \\ dy &= \frac{4t^2 - 2}{\sqrt{t^2 - 1}} dt. \end{aligned} \quad (6.4)$$

Both (6.3) and (6.4) are correct answers to the problem, (6.3) being the answer in terms of the variable x and (6.4) that in terms of t .

EXAMPLE 4. Find $D_x y$ if $y = 2t^3 - 1$, $x = t^2 + 1$.

SOLUTION. This type of problem was solved in Chapter 3 by dividing Dy by $D_t x$. In terms of differentials the same method can be written

$$dy = 6t^2 dt, \quad dx = 2t dt,$$

and

$$D_x y = \frac{dy}{dx} = \frac{6t^2 dt}{2t dt} = 3t.$$

This application of differentials is easily generalized. See Exercise 11, below.

EXAMPLE 5. Find $D_x y$ from the equation

$$2y^2 - 3xy + x^3 + 1/x = 12.$$

SOLUTION. This type of problem was solved in Chapter 3. Here we express the steps of the solution in terms of differentials. Applying the formulas for differentials to the separate terms,

$$4y \, dy - 3x \, dy - 3y \, dx + 3x^2 \, dx - (1/x^2) \, dx = 0.$$

Transposing terms in dx to the right-hand side

$$4y \, dy - 3x \, dy = 3y \, dx - 3x^2 \, dx + (1/x^2) \, dx.$$

Factoring,

$$(4y - 3x) \, dy = (3y - 3x^2 + 1/x^2) \, dx.$$

Dividing by dx and by $4y - 3x$,

$$\begin{aligned} D_x y &= \frac{dy}{dx} = \frac{3y - 3x^2 + 1/x^2}{4y - 3x} \\ &= \frac{3x^2y - 3x^4 + 1}{(4y - 3x)x^2}. \end{aligned}$$

EXERCISES

Find dy for each of the functions in Exercises 1-12.

1. $y = 4x^3 - 5x^2 + 2x + 11.$

2. $y = 3/x^2.$

3. $y = (2x^2 - 1)^4.$

4. $y = \sqrt{4 - x^2}.$

5. $y = (3x - 1)/(x^2 + 1).$

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

7. $y = \log(x^2 - 2) + \sin x.$

8. $y = \tan^{-1} x^2.$

9. $y = u^3 - 1, u = \sqrt{x - 1}.$ (Two answers.)

10. $y = a \sin u, u = e^x.$ (Two answers.)

11. Generalize the method of Example 4 by proving that if $y = f(t), x = g(t)$, then $D_x y = dy/dx$, wherein both differentials are determined with t as the independent variable.

Find $D_x y$ in Exercises 12-15 by the method of Exercise 11.

12. $y = a(1 - \cos t), x = a(t - \sin t).$

13. $y = 1/t - 2t^2, x = 2\sqrt{t}.$

14. $y = 1/u, x = u^2 + 4u - 10.$

15. $y = (u^2 - 1)/(u^2 + 1), x = 2u/(u^2 + 1).$

Find $D_x y$ in Exercises 16–20 by the method of Example 5.

16. $4x^2 + 9y^2 = 36$.

17. $4x^2y - 6x^2 + 3y^2 - 2y + 10 = 0$.

18. $xy^2 - 9(x + 4) = 0$.

19. $\sqrt{x^2 + y^2} - 3xy = 20$.

20. $e^{x^2+y^2} + \tan^{-1}(y/x) = 1$.

7. The Mean Value Theorem for Derivatives. We shall now take up an exceedingly useful theorem that is directly connected with differences, difference quotients, and differentials. By its means we can estimate the errors in approximations by differentials. But the greatest service of this

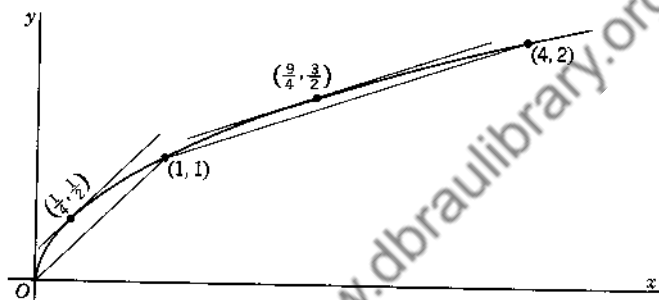


FIG. 5

theorem is its role in proving many important and practical results. One of these we shall take up in Art. 11. Others will be found in later chapters.

This theorem is called the *mean value theorem for derivatives* (to distinguish it from the *mean value theorem for integrals*), or the *theorem of the mean for derivatives*. We shall first illustrate its meaning graphically and then proceed to its proof.

With the function $y = \sqrt{x}$, graphed in Fig. 5, the mean value theorem tells us that if we connect 2 points of the curve by a chord, the points $(0, 0)$ and $(1, 1)$ for example, there is a point of the curve for which x lies between 0 and 1, where the tangent has the same slope as the chord. This is the point $(\frac{1}{4}, \frac{1}{2})$, as may be shown by calculating the slope of the chord and the slope of the tangent at this point:

$$(\text{slope of the chord}) = (1 - 0)/(1 - 0) = 1;$$

Since $D_x y = 1/(2\sqrt{x})$, $(\text{slope of the tangent at } x = \frac{1}{4}) = 1/(2\sqrt{1/4}) = 1$.

Similarly, if we draw a chord connecting $(1, 1)$ with $(4, 2)$, its slope is $(2 - 1)/(4 - 1) = \frac{1}{3}$. The tangent has the slope $\frac{1}{3}$ at the intermediate point $(\frac{9}{4}, \frac{3}{2})$, as may be readily verified.

The mean value theorem tells us that the fact illustrated above is true for every function $y = f(x)$ in an interval x_1 to x_2 in which the function is differentiable. The exact conditions of the theorem will be stated later. This theorem seems intuitively obvious from graphical considerations, but intuitive acceptability is not a substitute for logical proof. We first consider a special case, known as Rolle's theorem.

8. Rolle's Theorem. If a function $\phi(x)$ is continuous in the closed interval $x_1 \leq x \leq x_2$, and possesses a unique derivative in the open interval $x_1 < x < x_2$, and if $\phi(x_1) = \phi(x_2) = 0$, then there exists at least one value of x , say x' , in the interval, for which $\phi'(x') = 0$.

This theorem is illustrated graphically by Fig. 6. The graph of $y = \phi(x)$ intersects the x -axis at x_1 and x_2 . It is continuous and has a

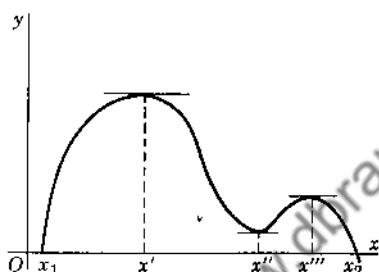


FIG. 6

single-valued slope at every point in the interval from x_1 to x_2 . Clearly, it must have its slope equal to 0 at some point in the interval. In this example there are three such points, x' , x'' , and x''' .

The proof of this theorem follows immediately from two theorems discussed earlier. We consider two cases. First, if $\phi(x)$ is everywhere 0, then $\phi'(x)$ is also everywhere 0, and the theorem holds. Second, if $\phi(x)$ is positive at some point in the interior of the interval, then by Theorem 6, Art. 11, Chapter 2, it must have a maximum in the interval. If this occurs at $x = x'$, then by Theorem 3, Art. 3, Chapter 5, $\phi'(x') = 0$. The argument is the same if $\phi(x)$ is negative at some point, and consequently has a minimum in the interval.

9. Proof of the Mean Value Theorem. We now proceed to the proof of the mean value theorem itself. The statement of this theorem is the following.

MEAN VALUE THEOREM. If $f(x)$ is continuous in the closed interval $x_1 \leq x \leq x_2$ and possesses a unique derivative at every point of the open interval $x_1 < x < x_2$, then there is at least one value of x , say $x = x'$, for which $x_1 < x' < x_2$, such that $\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x')$.

Figure 7 shows the graphical meaning of this theorem. The figure also suggests a way by which we can obtain a proof of the mean value theorem from Rolle's theorem. We write the equation of the chord PQ , which we represent temporarily by $y = \lambda(x)$ (Greek lambda). We now let $\phi(x) = f(x) - \lambda(x)$. Since, at x_1 and x_2 , $f(x)$ and $\lambda(x)$ are equal, we must have $\phi(x_1) = \phi(x_2) = 0$. Therefore, $\phi(x)$ satisfies the conditions for

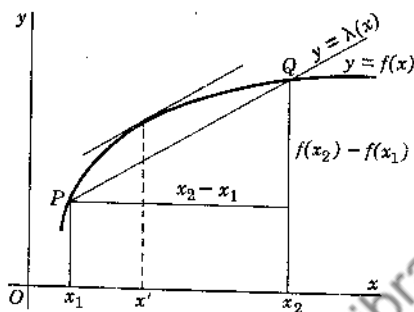


FIG. 7

Rolle's theorem, and $\phi'(x') = 0$ for some value x' between x_1 and x_2 . From this fact the mean value theorem will follow immediately. Let us fill in the steps.

Using the two-point form of the straight-line equation, we find the equation of the line through P and Q to be

$$y - f(x_1) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} (x - x_1),$$

or

$$y = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} (x - x_1).$$

The expression on the right-hand side is $\lambda(x)$. Subtracting $\lambda(x)$ from $f(x)$ to form $\phi(x)$,

$$\phi(x) = f(x) - f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1} (x - x_1). \quad (9.1)$$

From the way that $\phi(x)$ was formed we know that $\phi(x_1) = \phi(x_2) = 0$. This fact is easily verified from (9.1). Now, by Rolle's theorem, $\phi(x) = 0$ for some intermediate value, $x = x'$. Differentiating,

$$\phi'(x) = f'(x) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Setting $\phi'(x') = 0$,

$$f'(x') = \frac{f(x_2) - f(x_1)}{x_2 - x_1}, \quad (9.2)$$

which is the statement we set out to prove.

The mean value theorem can be stated in several forms, each of which has advantages for special purposes. In the form of equation (9.2) it emphasizes the fact that the difference quotient, formed for two values of x , is equal to the derivative at an intermediate value of x . It is common to express x' by the sum $x_1 + \theta(x_2 - x_1)$, where $0 < \theta < 1$. The formula now reads

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_1 + \theta \cdot \overline{x_2 - x_1}). \quad (9.3)$$

If we let $h = x_2 - x_1$,

$$\frac{f(x_1 + h) - f(x_1)}{h} = f'(x_1 + \theta h) \quad (9.4)$$

or

$$f(x_1 + h) = f(x_1) + hf'(x_1 + \theta h). \quad (9.5)$$

Since this equation holds for any values of x_1 and h as long as the conditions of the theorem are satisfied in the interval from x_1 to $x_1 + h$, we can drop the subscript and write

$$f(x + h) = f(x) + hf'(x + \theta h). \quad (9.6)$$

Writing $\Delta x = h$ and $\Delta f(x) = f(x + h) - f(x)$, we also have

$$\Delta f(x) = f'(x + \theta \cdot \Delta x) \Delta x. \quad (9.7)$$

The forms (9.5) and (9.6) suggest the use of the mean value theorem to determine a function at a point $x + h$, starting with the knowledge of the function at x . Since we cannot expect to know the precise value of $f'(x + \theta h)$, the best we could hope for would be an approximation of $f(x + h)$. The usefulness of the mean value theorem in connection with approximations will be discussed in Art. 10.

EXAMPLE 1. Find a value of x' satisfying the mean value theorem, Eq. 9.2, if $f(x) = x^{\frac{3}{2}}$, $x_1 = 0$, and $x_2 = 8$.

SOLUTION. We have $\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{8^{\frac{3}{2}} - 0^{\frac{3}{2}}}{8 - 0} = \frac{4}{8} = \frac{1}{2}$

Since

$$f'(x) = \frac{2}{3x^{\frac{1}{2}}},$$

we must have

$$\frac{2}{3(x')^{\frac{1}{2}}} = \frac{1}{2}.$$

Solving,

$$(x')^{\frac{1}{2}} = \frac{4}{3} \quad \text{and} \quad x' = \frac{64}{27} = 2\frac{10}{27}.$$

EXAMPLE 2. Show that if $0 < h < 1$

$$\frac{2h}{1-h^2} > \log \frac{1+h}{1-h} > 2h.$$

SOLUTION. Letting $f(x) = \log \frac{1+x}{1-x}$,

$$f'(x) = \frac{2}{1-x^2} \quad \text{and} \quad f'(x+\theta h) = \frac{2}{1-(x+\theta h)^2}.$$

From the mean value theorem, Eq. 9.6,

$$\log \frac{1+(x+h)}{1-(x+h)} = \log \frac{1+x}{1-x} + \frac{2h}{1-(x+\theta h)^2}, \quad 0 < \theta < 1.$$

Now, let $x = 0$,

$$\log \frac{1+h}{1-h} = 0 + \frac{2h}{1-\theta^2 h^2}.$$

If we were to make $\theta = 1$, the fraction $\frac{2h}{1-\theta^2 h^2}$ would be increased.

Hence,

$$\log \frac{1+h}{1-h} < \frac{2h}{1-h^2}.$$

If we were to make $\theta = 0$, the fraction $\frac{2h}{1-\theta^2 h^2}$ would be decreased.

Hence,

$$\log \frac{1+h}{1-h} > 2h.$$

Therefore,

$$\frac{2h}{1-h^2} > \log \frac{1+h}{1-h} > 2h.$$

EXERCISES

In each of Exercises 1-5 find a value of x' satisfying the mean value theorem, Eq. 9.2.

1. $f(x) = x^2$;

$x_1 = 0, x_2 = 4.$

2. $f(x) = 1/x$;

$x_1 = 1, x_2 = 4.$

3. $f(x) = \log x$;

$x_1 = 1, x_2 = e.$

4. $f(x) = \tan x$;

$x_1 = 0, x_2 = \pi/4.$

5. $f(x) = e^{-x}$;

$x_1 = 0, x_2 = 1.$

6. If $f(x) = x^{2/3}$, $x_1 = -1$, and $x_2 = 1$, there is no value of x' between x_1 and x_2 for which (9.2) holds. Why? If $x_1 = -1$ and $x_2 = 27$, there is such a value of x' . Does this fact contradict the mean value theorem? Explain.

7. Show that

$$\frac{h}{1+h} < \log(1+h) < h, \quad \text{if } 0 < h.$$

8. Show that

$$\frac{h}{2\sqrt{1+h}} < \sqrt{1+h} - 1 < \frac{1}{2}h, \quad \text{if } 0 < h.$$

9. Verify Rolle's theorem for the function $f(x) = \sin x - \frac{1}{2}x$ in an interval one end of which is $x = 0$.

10. Verify Rolle's theorem for the function

$$\begin{aligned} f(x) &= x \log x, & \text{for } x > 0, \\ &= 0, & \text{for } x = 0, \end{aligned}$$

in the interval $0 \leq x \leq 1$.

10. The Mean Value Theorem and Differential Approximations. In Art. 3 we found approximations to functional values in the neighborhood of known values by means of the approximate equation

$$y + \Delta y \approx y + dy = f(x) + f'(x) dx.$$

In order to compare this approximate equation with the equation expressing the mean value theorem, we note that

$$y + \Delta y = f(x + h), \quad dx = h.$$

The approximate equation then reads

$$f(x + h) \approx f(x) + hf'(x).$$

The mean value theorem, on the other hand, is

$$f(x + h) = f(x) + hf'(x + \theta h), \quad 0 < \theta < 1.$$

The error incurred in using the differential approximation is thus the difference between $hf'(x)$ and $hf'(x + \theta h)$:

$$|\text{error}| = |hf'(x) - hf'(x + \theta h)|.$$

We can hardly expect to compute the exact value of the error. It is often possible, however, to set an upper bound to the error, and thus show that the approximation lies within an acceptable range. As a first example we return to Example 2, Art. 3.

EXAMPLE 1. Estimate the accuracy of the error and relative error approximated in Example 2, Art. 3.

SOLUTION. In this example we found

$$|dA| = 2.4.$$

Now

$$\begin{aligned} |\Delta A| &= 2(12 \pm .1\theta)(.1), \text{ by Eq. 9.7} \\ &= 2.4 \pm .02\theta, \quad \text{where } 0 < \theta < 1. \end{aligned}$$

Our approximation by dA is therefore in error by not more than $.02\theta$, which is less than $.02$, and hence is accurate to at least one decimal place.

In estimating the relative error, we used

$$|dA/A| = .0166 \dots$$

We know, now, that

$$\left| \frac{\Delta A}{A} \right| = \frac{2.4 \pm .02\theta}{144} = .0166 \dots \pm .00014 \dots$$

and hence the estimate of relative error was correct to 3 decimal places.

EXAMPLE 2. Estimate the error in the approximation to $\sin 31^\circ$ in Example 3, Art. 3.

SOLUTION. In Example 3 we used

$$\begin{aligned} dy &= \cos x \, dx \\ &= \cos \frac{\pi}{6} (.01745) \\ &= (.86603)(.01745) = .01511. \end{aligned}$$

Now, by Eq. 9.7,

$$\Delta y = \cos(\pi/6 + .01745\theta) \cdot .01745, \quad \text{where } 0 < \theta < 1.$$

The error in the approximation is therefore

$$\Delta y - dy = [\cos(\pi/6 + .01745\theta) - \cos \pi/6] \cdot .01745.$$

By the trigonometric formula,

$$\cos A - \cos B = -2 \sin \frac{1}{2}(A + B) \cdot \sin \frac{1}{2}(A - B),$$

we have

$$\Delta y - dy = [-2 \sin(\pi/6 + .01745\theta/2) \cdot \sin .01745\theta/2] \cdot .01745.$$

Now, we know that $\sin \pi/6 = .5$, and that $\sin(\pi/6 + .01745\theta/2) < .6$, since $\sin x$ increases less rapidly than x itself. Moreover,

$$\sin .01745\theta/2 < .01745\theta/2 < .00873, \quad \text{since } \theta < 1.$$

Hence

$$|\Delta y - dy| < 2(.6)(.00873)(.01745) < .000183.$$

We conclude that the approximate value of $\sin 31^\circ$, which was calculated as .51511, is too great, since $\Delta y - dy$ is negative, by an amount not exceeding .000183.

EXERCISES

1. Show that the value of $1/(1.01)^2$ found in Example 1, Art. 3, is too small, but that the error is less than .001.

2. Show that $\sqrt[3]{17} = 4\frac{1}{8} - E$, where $E < 1/264$. [Hint: In estimating the error use the fact that $\sqrt[3]{17} < \frac{25}{8}$.]

3. Show that $\sqrt[3]{8.05} = 2.004166\dots - E$, where $E < .000084$. [Hint: $\sqrt[3]{8 + .050} < 2.02$, since $(2.02)^3 > 2 \cdot 2 \cdot (2.02) = 8.08$.]

4. Show that $.88240 < \cos 28^\circ < .88348$.

5. Given $\log 50 = 3.911923$, (a) show that $3.915907 < \log 50.2 < 3.915923$. (b) Approximate the value of $\log 50.02$ by means of a differential and estimate the error.

11. The Fundamental Theorem on Primitive Functions. With the help of the mean value theorem we can prove a fundamental fact about primitive functions that we have assumed until now.

THEOREM. Let $F(x)$ and $G(x)$ be two functions having the same derived function. That is, $F'(x) = G'(x)$. Then

$$F(x) - G(x) = \text{a constant.}$$

In order to prove this let us set $F(x) - G(x) = \phi(x)$. Then, by hypothesis, for any value of x ,

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

Comparing the values of $\phi(x)$ at any two points x_1 and x_2 , we have, by the mean value theorem,

$$\phi(x_2) - \phi(x_1) = (x_2 - x_1)\phi'(x'),$$

where x' lies between x_1 and x_2 . Since $\phi'(x') = 0$, it follows that $\phi(x_2) = \phi(x_1)$. Since this is true for any values of x_1 and x_2 , $\phi(x)$ always has the same value; i.e.,

$$\phi(x) = F(x) - G(x) = \text{a constant.}$$

In words, this theorem says that every primitive function of a given derived function can be obtained from any particular primitive function by the addition of a constant. We now have complete justification for expressing the general form of a primitive function of any given function as the sum of any of its inverse derivations and an arbitrary constant.

Technique of Integration

1. Introduction. The calculation of derived functions and the uses to which they were put suggested the opposite problem of passing from the derived function to its primitives. We did this in Chapters 6, 10, and 11 by simply inverting the formulas for derived functions. Thus far, however, the process of inverse differentiation stands in sharp contrast to differentiation. For we have formulas for differentiating almost any conceivable combination of algebraic, trigonometric, logarithmic, and exponential functions, whereas we can apply inverse differentiation successfully in only a relatively small number of cases. Compare the following, for example.

$$\begin{array}{ll}
 D\sqrt{1-x^2} & \text{and} \quad D^{-1}\sqrt{1-x^2}; \\
 De^{ax} \cos bx & \text{and} \quad D^{-1}e^{ax} \cos bx; \\
 D\sqrt{\sin x} & \text{and} \quad D^{-1}\sqrt{\sin x}; \\
 D\sqrt{x} \sin x & \text{and} \quad D^{-1}\sqrt{x} \sin x.
 \end{array}$$

All the above differentiations can be carried out with ease. The first two inverse differentiations are beyond our present technique, and the third and fourth cannot be performed at all.

The formulas for differentiation were developed in a step-by-step fashion, by applying the definition of the derivative first to the simplest functions, and continuing in the same way to the last functions introduced. At the same time rules were derived for dealing with combinations of functions, such as sums, products, quotients, powers, roots, and, most important of all, a function of a function, the so-called chain rule. Because of the latter rule we can differentiate such a fantastic function as

$$\sqrt{e^{\sin \{ \tan^{-1} [\log (1+x^2)^{1/3}] \}}}$$

Our formulas for inverse differentiation were found in a different way. We obtain an inverse differentiation formula only by inverting a known formula for differentiation. We cannot find primitive functions of all sorts of combinations of functions, since it is not always possible to identify these combinations as the derivatives of known functions. Since

we can only find inverse derivatives of those functions that can be made to fit the inverted differentiation formulas, the technique of inverse differentiation consists in breaking up given functions, or transforming them, so that the formulas can be brought to bear upon them. The process is much like puzzle-solving.

In this chapter we shall first review the formulas found in earlier chapters, and then proceed to the study of special devices for giving these formulas the broadest possible application. Success in inverse differentiation depends to a considerable extent upon a facile knowledge of fundamental identities of algebra and trigonometry, as well as of the formulas for differentiation.

The close relationship between inverse differentiation and integration was brought out in Chapter 8, and the symbol $\int$ was introduced as a substitute for D^{-1} . Until now we have often emphasized the character of the operation by using D^{-1} . In a systematic study of inverse differentiation some advantages result from the use of the integral symbol, and we shall do so hereafter. For the most part we shall also replace the symbol D by d/dx , introduced in Chapter 12. For conciseness, we shall use the term "integration" in place of "inverse differentiation" much of the time. We call $\int f(x) dx$ the "integral of $f(x)$ with respect to x ." The function that follows the integral sign, $f(x)$, is called the *integrand*.

2. Change of Variable. Each of the basic formulas that was derived in earlier chapters was generalized at the time it was introduced. Thus, for example, the formula

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C$$

was generalized to

$$\int u^n \left(\frac{du}{dx} \right) dx = \frac{u^{n+1}}{n+1} + C,$$

where u is a function of x . The justification of these more general formulas rests on the fact that they can be verified by differentiation. Thus,

$$\frac{d}{dx} \int u^n \left(\frac{du}{dx} \right) dx = u^n \left(\frac{du}{dx} \right),$$

since the derivative of an inverse derivative, or integral, is the integrand, provided that the differentiation and inverse differentiation are with respect to the same variable. Differentiating the right-hand side, we obtain the same result.

$$\frac{d}{dx} \left[\frac{u^{n+1}}{n+1} + C \right] = \frac{d}{du} \left[\frac{u^{n+1}}{n+1} + C \right] \frac{du}{dx} = u^n \left(\frac{du}{dx} \right).$$

These generalized formulas can be written more concisely by the use of differentials. Thus, since $(du/dx) dx = du$, we can write

$$\int u^n (du/dx) dx = \int u^n du.$$

This form suggests that when we introduce $u =$ a function of x in the integral, we are actually changing our point of view with regard to the independent variable. It was x initially, and we now change it to u . After performing the integration we can change it back to x . By this transformation, replacing x by a function of u , or u by a function of x , $\int g(x) dx$ becomes $\int f(u) du$, or vice versa. We can formalize the steps for finding one of these integrals from the other by means of differentials, leading to formula 2.1 below.

It is often convenient to think of the problem of integration as that of finding a function having a given differential. Thus, when we write $\int 2x dx = x^2 + C$, we are saying that $x^2 + C$ is a function whose differential is $2x dx$. Now, according to the facts brought out in Art. 5, Chapter 12, if

$$dy/du = f(u), \quad \text{then} \quad dy = f(u) du,$$

and this statement is correct no matter what the independent variable may be. If u is independent, it follows from the definition of a differential. If u is dependent upon some other variable, say $u = \phi(x)$, it is still true that $dy = f(u) du$, where $du = \phi'(x) dx$. Since $f(u) du = f[\phi(x)]\phi'(x) dx$, it follows that a function whose differential with respect to u is $f(u) du$ is, at the same time, a function whose differential with respect to x is $f[\phi(x)]\phi'(x) dx$. That is,

$$\int f(u) du = \int f[\phi(x)]\phi'(x) dx. \quad (2.1)$$

We might expect the result of $\int f(u) du$ to be expressed in terms of u , and that of $\int f[\phi(x)]\phi'(x) dx$ in terms of x , but these two results must be identical when u is replaced by $\phi(x)$, except for the addition of a constant. Equation 2.1 is a formula for change of variable in an integral. As an exercise, the student should verify this formula by differentiating both sides with respect to x .

Each of the inverse differentiation formulas listed in Chapters 6, 10, and 11 can be generalized by the use of formula 2.1. Thus, for example,

$$\int u^n du = u^{n+1}/(n+1) + C,$$

$$\int \sin u du = -\cos u + C, \text{ etc.}$$

where u is any differentiable function of x .

The most obvious use of these general formulas is that of replacing a complicated looking integral with respect to x by an integral with respect to u that fits one of the formulas. For example,

$$\begin{aligned}\int (x^2 + 1)^n 2x \, dx &= \int u^n \, du, \quad \text{where } u = x^2 + 1, \\ &= u^{n+1}/(n+1) + C = (x^2 + 1)^{n+1}/(n+1) + C.\end{aligned}$$

Here, both the change of variable and the resulting integral are suggested by inspection. The same is true, although the substitution is slightly less obvious, in the example:

$$\begin{aligned}\int \frac{1}{x^2} \tan \frac{1}{x} \, dx &= - \int \tan u \, du, \quad \text{where } u = 1/x, \\ &= - \log |\sec u| + C = \log |\cos u| + C \\ &= \log |\cos 1/x| + C.\end{aligned}$$

Comparing these changes of variable with the general transformation formula 2.1, we see that we have found by inspection the functions $\phi(x)$, $\phi'(x)$, and $f[\phi(x)]$. Thus, in the second example,

$$\phi(x) = 1/x, \quad \phi'(x) = -1/x^2, \quad \text{and} \quad f[\phi(x)] = f(u) = \tan u.$$

It is not always easy to identify the three essential parts of an integrand, $\phi(x) = u$, $\phi'(x) \, dx = du$, and $f[\phi(x)] = f(u)$, that will transform a complicated integral with respect to x into a standard form of an integral with respect to u . A substitution may be suggested by the presence of some particular expression in the integrand, and it is then desirable to carry through the substitution to discover the form of the resulting integral. In such cases the substitution may be in the form $x = \psi(u)$. We can then employ the transformation formula

$$\int g(x) \, dx = \int g[\psi(u)] \psi'(u) \, du, \quad (2.2)$$

which expresses the same relation as (2.1), with x and u interchanged. Thus, suppose we try to transform the integral $\int \sqrt{1-x^2} \, dx$ by putting $x = \sin u = \psi(u)$. We have, by (2.2),

$$\begin{aligned}\int \sqrt{1-x^2} \, dx &= \int \sqrt{1-\sin^2 u} (\cos u) \, du \\ &= \int \cos^2 u \, du.\end{aligned}$$

The integral $\int \cos^2 u \, du$ will be discussed in Art. 8, and transformations of the type here illustrated will be taken up in Art. 9.

If a change of variable is employed in order to find a primitive function for the purpose of evaluating a definite integral, the final step of changing the result back into an expression in the original variable can

be avoided, often advantageously. For example, suppose we wish to find $\int_0^1 \sqrt{1+x^2} 2x dx$. Putting $u = 1 + x^2$, this becomes $\int_1^2 u^{\frac{1}{2}} du$, where the new limits of integration are obtained as follows: since $u = 1 + x^2$, when $x = 0$, $u = 1$, and when $x = 1$, $u = 2$.

3. Integration Formulas. The formulas for integration that were derived in earlier chapters are listed below for reference and review. Formulas *I-22-I-24* will be derived in this chapter, *I-22* in Art. 6, and *I-23* and *I-24* in Art. 9. The formulas are expressed in terms of a variable u that is to be thought of as representing the independent variable itself, or any differentiable function of that variable. It is these formulas to which we shall try to fit all our integration problems.

$$\int u^n du = u^{n+1}/(n+1) + C, \text{ if } n \neq -1. \quad (I-1)$$

$$\int u^{-1} du = \int \frac{1}{u} du = \log |u| + C. \quad (I-14)$$

$$\int \cos u du = \sin u + C. \quad (I-2)$$

$$\int \sin u du = -\cos u + C. \quad (I-3)$$

$$\int \sec^2 u du = \tan u + C. \quad (I-4)$$

$$\int \csc^2 u du = -\cot u + C. \quad (I-5)$$

$$\int \sec u \tan u du = \sec u + C. \quad (I-6)$$

$$\int \csc u \cot u du = -\csc u + C. \quad (I-7)$$

$$\int \tan u du = \log |\sec u| + C. \quad (I-19)$$

$$\int \cot u du = \log |\sin u| + C. \quad (I-18)$$

$$\int \sec u du = \log |\sec u + \tan u| + C. \quad (I-20)$$

$$\int \csc u du = \log |\csc u - \cot u| + C. \quad (I-21)$$

$$\int \frac{1}{\sqrt{a^2 - u^2}} du = \sin^{-1} \frac{u}{a} + C, \text{ if } |u| < |a| \quad (I-8)$$

$$= -\cos^{-1} \frac{u}{a} + C, \text{ if } |u| < |a|. \quad (I-9)$$

$$\int \frac{1}{a^2 + u^2} du = \frac{1}{a} \tan^{-1} \frac{u}{a} + C. \quad (I-10)$$

$$= -\frac{1}{a} \cot^{-1} \frac{u}{a} + C. \quad (I-11)$$

$$\int \left| \frac{1}{u\sqrt{u^2 - a^2}} \right| du = \frac{1}{a} \sec^{-1} \frac{u}{a} + C, \text{ if } |u| > |a| \quad (I-12)$$

$$= -\frac{1}{a} \csc^{-1} \frac{u}{a} + C, \text{ if } |u| > |a|. \quad (I-13)$$

$$\int \frac{1}{a^2 - u^2} du = \frac{1}{2a} \log \left| \frac{u+a}{u-a} \right| + C. \quad (I-22)$$

$$= \frac{1}{2a} \log \frac{u+a}{u-a} + C, \text{ if } |u| > |a|. \quad (I-22a)$$

$$= -\frac{1}{2a} \log \frac{a+u}{a-u} + C, \text{ if } |u| < |a|. \quad (I-22b)$$

$$\int \frac{1}{\sqrt{u^2 + a^2}} du = \log(u + \sqrt{u^2 + a^2}) + C. \quad (I-23)$$

$$\int \frac{1}{\sqrt{u^2 - a^2}} du = \log(u + \sqrt{u^2 - a^2}) + C, \text{ if } |u| > |a|. \quad (I-24)$$

$$\int e^u du = e^u + C. \quad (I-16)$$

$$\int b^u du = b^u \log_e b + C. \quad (I-17)$$

In addition to these formulas, we have two important theorems.

1. $\int [u(x) + v(x)] dx = \int u(x) dx + \int v(x) dx.$
2. $\int a u(x) dx = a \int u(x) dx.$

Theorem 1 can be extended to sums and differences of any number of functions. We have no rule for integrating the product of two functions. A method that is often useful in connection with products will be discussed in Art. 7.

The integrals in the following exercises are types that have been taken up in earlier chapters, or that can be reduced to those types by simple changes in the variables. They will serve to review the use of the formulas listed above, as well as to give practice in substituting a new variable.

EXERCISES

Calculate the integrals in Exercises 1-32.

$$1. \int (15x^2 - 4x + \frac{1}{3}) dx = 5x^3 - 2x^2 + \frac{1}{3}x + C.$$

$$2. \int \frac{3}{x^2} dx.$$

$$3. \int \frac{3}{x} dx = 3 \log |x| + C.$$

$$4. \int \left(\sqrt{x} - \frac{1}{\sqrt{x}} \right) dx.$$

$$5. \int 2\sqrt{x} \, dx = \frac{4}{3}x^{\frac{3}{2}} + C.$$

$$6. \int \left(\frac{2}{x^3} - \frac{1}{x} + 2x \right) dx.$$

$$7. \int x\sqrt{1+x^2} \, dx = \frac{1}{3}(1+x^2)^{\frac{3}{2}} + C.$$

$$8. \int \frac{x \, dx}{\sqrt{1+x^2}}.$$

$$9. \int \frac{x \, dx}{1+x^2} = \log \sqrt{1+x^2} + C.$$

$$10. \int \sin 2x \, dx.$$

$$11. \int \sec^2 \frac{x}{2} \, dx = 2 \tan \frac{x}{2} + C.$$

$$12. \int \sec^2 2x \, dx.$$

$$13. \int x \csc x^2 \cot x^2 \, dx = -\frac{1}{2} \csc x^2 + C.$$

$$14. \int \cos x e^{\sin x} \, dx.$$

$$15. \int 10^x \, dx = 10^x \log_{10} e + C.$$

$$16. \int \frac{\log |x|}{x} \, dx.$$

$$17. \int \frac{dx}{\sqrt{4-x^2}} = \sin^{-1} \frac{x}{2} + C.$$

$$18. \int \frac{2 \, dx}{4+x^2}.$$

$$19. \int \tan 3x \, dx = \frac{1}{3} \log |\sec 3x| + C.$$

$$20. \int \csc \frac{x}{3} \, dx.$$

$$21. \int e^{\tan^{-1} x} / (1+x^2) \, dx = e^{\tan^{-1} x} + C.$$

$$22. \int \cot x \log |\sin x| \, dx.$$

$$23. \int x^3(x^4+9)^{\frac{1}{2}} \, dx = \frac{1}{10}(x^4+9)^{\frac{3}{2}} + C.$$

$$24. \int \frac{x \, dx}{(x^2-9)^{\frac{3}{2}}}.$$

$$25. \int \frac{x \, dx}{\sqrt{9-x^4}} = \frac{1}{2} \sin^{-1} \frac{x^2}{3} + C.$$

$$26. \int \frac{x \, dx}{9+x^4}.$$

$$27. \int \frac{\sin^{-1} x}{\sqrt{1-x^2}} \, dx = \frac{1}{2} (\sin^{-1} x)^2 + C.$$

$$28. \int (x-1)e^{x^2-2x} \, dx.$$

$$29. \int \frac{e^x \log |e^x - 1|}{e^x - 1} \, dx = \frac{1}{2} [\log |e^x - 1|]^2 + C.$$

$$30. \int x \sin(x^2 + 2) \, dx.$$

$$31. \int \tan x \log |\cos x| \, dx = -\frac{1}{2} (\log |\cos x|)^2 + C.$$

$$32. \int \frac{x \tan \sqrt{x^2-1}}{\sqrt{x^2-1}} \, dx.$$

Calculate the definite integrals in Exercises 33–40.

$$33. \int_0^4 3x\sqrt{x^3+9} \, dx = 98.$$

$$34. \int_0^3 \frac{x \, dx}{x^2+16} = \log \frac{5}{4}.$$

$$35. \int_0^{\pi/2} \sin^2 x \cos x \, dx = \frac{1}{3}.$$

$$36. \int_0^{\pi/4} \sec^2 x \tan x \, dx = \frac{1}{2}.$$

$$37. \int_0^1 \frac{2 \tan^{-1} x}{1+x^2} \, dx = \frac{\pi^2}{16}.$$

$$38. \int_2^3 \frac{2x+6}{x^2+6x-9} \, dx = \log \frac{16}{7}.$$

$$39. \int_{-\pi/4}^0 \sec^2 x \cdot e^{\tan x} \, dx = 1 - 1/e.$$

$$40. \int_{-1}^1 \frac{dx}{\sqrt{1-x^2} \cdot \sin^{-1} x} = \log 3.$$

4. A Quadratic Reduction. A simple transformation of a quadratic expression will change many fractions into forms found in the formulas for integration. Let us consider the following examples.

EXAMPLE 1. $\int \frac{1}{x^2 + 2x + 5} dx.$

SOLUTION. By writing

$$\begin{aligned} x^2 + 2x + 5 &= x^2 + 2x + 1 + 4 \\ &= (x + 1)^2 + 4, \end{aligned}$$

we can express the above integral in the form

$$\int \frac{1}{u^2 + 4} du, \text{ where } u = x + 1.$$

Hence, using formula I-10,

$$\begin{aligned} \int \frac{1}{x^2 + 2x + 5} dx &= \int \frac{1}{u^2 + 4} du \\ &= \frac{1}{2} \tan^{-1} \frac{u}{2} + C \\ &= \frac{1}{2} \tan^{-1} \frac{x + 1}{2} + C. \end{aligned}$$

EXAMPLE 2. $\int \frac{1}{\sqrt{4x - x^2 - 3}} dx.$

SOLUTION. We write $4x - x^2 - 3 = -(x^2 - 4x + 3)$. By adding 4 to the first 2 terms and subtracting 4 from the third term to balance this addition, $x^2 - 4x + 3$ becomes

$$x^2 - 4x + 4 - 1 = (x - 2)^2 - 1$$

and

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

Hence,

$$\begin{aligned} \int \frac{1}{\sqrt{4x - x^2 - 3}} dx &= \int \frac{1}{\sqrt{1 - (x - 2)^2}} dx \\ &= \int \frac{1}{\sqrt{1 - u^2}} du, \text{ where } u = x - 2, \\ &= \sin^{-1} u + C, \text{ by (I-8),} \\ &= \sin^{-1} (x - 2) + C. \end{aligned}$$

EXAMPLE 3. $\int \frac{1}{\sqrt{21 - 18x - 2x^2}} dx.$

SOLUTION. We write $21 - 18x - 2x^2 = -2(x^2 + 9x - 21/2)$. By adding $\frac{81}{4}$ to the first 2 terms to make a perfect square, and subtracting $\frac{81}{4}$ from the third term to balance this addition, $x^2 + 9x - \frac{21}{2}$ becomes

$$x^2 + 9x + \frac{81}{4} - \frac{123}{4} = \left(x + \frac{9}{2}\right)^2 - \left(\frac{\sqrt{123}}{2}\right)^2$$

and

$$-2\left(x^2 + 9x - \frac{21}{2}\right) = 2\left[\left(\frac{\sqrt{123}}{2}\right)^2 - \left(x + \frac{9}{2}\right)^2\right].$$

Hence,

$$\begin{aligned} \int \frac{1}{\sqrt{21 - 18x - 2x^2}} dx &= \int \frac{1}{\sqrt{2} \sqrt{\left(\frac{\sqrt{123}}{2}\right)^2 - \left(x + \frac{9}{2}\right)^2}} dx \\ &= \frac{1}{\sqrt{2}} \int \frac{1}{\sqrt{a^2 - u^2}} du \end{aligned}$$

where $a = \sqrt{123}/2$ and $u = x + \frac{9}{2}$. The required integral is therefore equal to $\frac{1}{\sqrt{2}} \sin^{-1} \frac{u}{a} + C$, or $\frac{1}{\sqrt{2}} \sin^{-1} \frac{(2x + 9)}{\sqrt{123}} = C$.

EXERCISES

Calculate the following integrals.

1. $\int \frac{dx}{x^2 - 10x + 29} = \frac{1}{2} \tan^{-1} \frac{x-5}{2} + C.$
2. $\int \frac{dx}{\sqrt{10x - 21 - x^2}}$
3. $\int \frac{dx}{\sqrt{12 - 8x - 4x^2}} = \frac{1}{2} \sin^{-1} \frac{x+1}{2} + C.$
4. $\int \frac{dx}{7x^2 + 28x + 203}$
5. $\int \frac{dx}{9x^2 - 12x + 5} = \frac{1}{3} \tan^{-1} (3x + 2) + C.$
6. $\int \frac{dx}{(x+2)\sqrt{x^2 + 4x + 3}}$
7. $\int \frac{(x-1) dx}{x^2 - 4x + 8} = \log \sqrt{x^2 - 4x + 8} + \frac{1}{2} \tan^{-1} \frac{x-2}{2} + C.$

[Hint: Write $\frac{x-1}{x^2 - 4x + 8} = \frac{x-2}{x^2 - 4x + 8} + \frac{1}{x^2 - 4x + 8}$].

8. $\int \frac{(x+1) dx}{\sqrt{12 - 4x - x^2}}$

$$9. \int \frac{8x + 12}{x^2 + 16x + 100} dx = 4 \log(x^2 + 16x + 100) - \frac{2}{3} \tan^{-1} \frac{x+8}{6} + C.$$

$$10. \int \frac{3x - 2}{\sqrt{4x - x^2}} dx.$$

5. Integration of Some Rational Fractions. The list of formulas in Art. 3 contains only a few that deal with rational fractions. We have (I-1) for negative n , (I-14), (I-10), (I-11), and (I-22). (I-22) is yet to be derived. This seems to place a narrow limitation on our ability to integrate fractions. We can extend our compass considerably in this respect by the method of the present article. We shall consider proper fractions* with polynomials in the numerators, and denominators consisting of any number of linear and quadratic factors. The fraction

$$\frac{18 + 11x - x^2}{(x-1)(x+1)(x^2+3x+3)}$$

is an example. The method will be that of breaking such a fraction into a sum of partial fractions. The fraction given above is equal to

$$\frac{2}{x-1} - \frac{3}{x+1} + \frac{x-3}{x^2+3x+3}, \text{ as we shall see later (Example 2).}$$

In finding the partial fractions for fractions of the type mentioned above we shall be guided by the following facts.† If the factors of the denominator are all different as in the above example, the fraction can be expressed as the sum of proper fractions, each of which has its denominator equal to one of the factors in the original denominator. For linear factors the numerators must be constants, since the numerator of a proper fraction is of lower degree than the denominator. For the same reason, the partial fractions having quadratic denominators have their numerators constants or linear expressions. For example,

$$\frac{3x+11}{(x+2)(x+3)} = \frac{A}{x+2} + \frac{B}{x+3},$$

and

$$\frac{x^2+11x-10}{(x+2)(x+3)(x^2+1)} = \frac{C}{x+2} + \frac{D}{x+3} + \frac{Ex+F}{x^2+1},$$

in which the capital letters represent constants. If a factor of the denominator is repeated n times, for example, $(x+2)^n$, the partial

* A rational fraction is "proper" if its numerator is of lower degree than its denominator.

† See *College Algebra*, by Fine, p. 241.

fractions may include fractions for each power of this factor from 1 to n , but if this factor is linear the numerators of all these fractions are constants, whereas if it is quadratic the numerators are all either linear or constants. For example,

$$\frac{2x^3 + 12x^2 + 19x + 5}{(x+2)^2(x^2+3x+1)} = \frac{A}{x+2} + \frac{B}{(x+2)^2} + \frac{Cx+D}{x^2+3x+1}$$

and

$$\frac{x^2-5}{(x^2+1)^3} = \frac{Mx+N}{x^2+1} + \frac{Px+Q}{(x^2+1)^2} + \frac{Rx+S}{(x^2+1)^3},$$

where the capital letters stand for constants.

Before taking up the process of breaking a fraction into partial fractions, we may observe that we can integrate all partial fractions of the type described, with one exception which will take further study. We consider them one by one.

1. Linear denominator to the first power.

$$\int \frac{A}{ax+b} dx = \frac{A}{a} \log |ax+b| + C, \text{ by (I-14).}$$

2. Linear denominator raised to the n th power, $n \neq 1$.

$$\int \frac{A}{(ax+b)^n} dx = \frac{A}{a} \frac{(ax+b)^{-n+1}}{-n+1} + C, \text{ by (I-1).}$$

3. Quadratic denominator to the first power. We shall calculate an integral of the type $\int \frac{Ax+B}{ax^2+bx+c} dx$, and it will be easy to see that the method is quite general.

Let us calculate

$$\int \frac{3x+2}{2x^2+5x+7} dx.$$

First, we take out the factor $\frac{3}{4}$, obtaining

$$\frac{3}{4} \int \frac{4x + \frac{8}{3}}{2x^2 + 5x + 7} dx.$$

This is done so that the linear term in the numerator will be the derivative of the quadratic term in the denominator. We now add and subtract $\frac{7}{3}$ in the numerator, obtaining

$$\frac{3}{4} \int \frac{4x + 5 - \frac{7}{3}}{2x^2 + 5x + 7} dx.$$

This makes the first two terms in the numerator the derivative of the denominator. We can now express the integral as the sum of two integrals.

$$\begin{aligned} \frac{3}{4} \int \frac{4x + 5 - \frac{7}{3}}{2x^2 + 5x + 7} dx &= \frac{3}{4} \left[\int \frac{(4x + 5) dx}{2x^2 + 5x + 7} - \int \frac{\frac{7}{3} dx}{2x^2 + 5x + 7} \right] \\ &= \frac{3}{4} \log |2x^2 + 5x + 7| - \frac{3}{4} \cdot \frac{7}{3} \cdot \frac{2}{\sqrt{31}} \tan^{-1} \frac{4x + 5}{\sqrt{31}} + C. \end{aligned}$$

The second fraction in the brackets was integrated by the method of Art. 4.

Fractions of the type $\frac{Cx + D}{(ax^2 + bx + c)^n}$, in which the denominator is a quadratic to a power higher than the first, are more troublesome. They do not arise in problems so frequently as to make their study imperative at this time. A method of dealing with such fractions will be shown in Art. 11.

6. Examples of Integration by Partial Fractions. The method of breaking a fraction into partial fractions, and integration based on this decomposition, is illustrated in the examples that follow.

EXAMPLE 1. Evaluate $\int \frac{3x + 11}{(x + 2)(x + 3)} dx$.

SOLUTION. We write

$$\begin{aligned} \frac{3x + 11}{(x + 2)(x + 3)} &= \frac{A}{x + 2} + \frac{B}{x + 3} \\ &= \frac{A(x + 3) + B(x + 2)}{(x + 2)(x + 3)}. \end{aligned}$$

Since these two fractions are merely different ways of writing the same quantity, and their denominators are identical, their numerators must be identically equal, that is, equal for all values of x .

$$A(x + 3) + B(x + 2) \equiv 3x + 11.$$

[$\equiv$ is the strong equality sign which symbolizes identity.]

Putting $x = -2$ so that B will disappear, $A = -6 + 11 = 5$.

Putting $x = -3$, so that A will disappear, $-B = -9 + 11 = 2$, $B = -2$. Hence,

$$\frac{3x + 11}{(x + 2)(x + 3)} = \frac{5}{x + 2} - \frac{2}{x + 3}.$$

and

$$\begin{aligned}\int \frac{3x+11}{(x+2)(x+3)} dx &= \int \frac{5}{x+2} dx - \int \frac{2}{x+3} dx \\ &= 5 \log |x+2| - 2 \log |x+3| + C.\end{aligned}$$

EXAMPLE 2. Evaluate $\int \frac{18+11x-x^2}{(x-1)(x+1)(x^2+3x+3)} dx$.

SOLUTION. We write

$$\begin{aligned}\frac{18+11x-x^2}{(x-1)(x+1)(x^2+3x+3)} &= \frac{A}{x-1} + \frac{B}{x+1} + \frac{Cx+D}{x^2+3x+3} \\ &= \frac{A(x+1)(x^2+3x+3) + B(x-1)(x^2+3x+3) + (Cx+D)(x-1)(x+1)}{(x-1)(x+1)(x^2+3x+3)}\end{aligned}$$

Again, as in Example 1, the two numerators must be identically equal.

$$\begin{aligned}A(x+1)(x^2+3x+3) + B(x-1)(x^2+3x+3) + (Cx+D)(x-1)(x+1) \\ \equiv 18+11x-x^2.\end{aligned}$$

The method of evaluating A and B in Example 1 is less useful here. This method is usually best only when all the factors of the denominator are linear and none are repeated. Instead, we multiply out and collect like powers of x .

$$\begin{aligned}(A+B+C)x^3 + (4A+2B+D)x^2 + (6A-C)x + (3A-3B-D) \\ \equiv 18+11x-x^2.\end{aligned}$$

Two polynomials that are identically equal have the coefficients of like powers of x equal. Hence,

$$\begin{aligned}A+B+C &= 0, \\ 4A+2B+D &= -1, \\ 6A-C &= 11, \\ 3A-3B-D &= 18.\end{aligned}$$

We now solve these equations simultaneously for A , B , C , and D . Adding the first and third, and the second and fourth,

$$\begin{aligned}7A+B &= 11, \\ 7A-B &= 17.\end{aligned}$$

From these two equations we find that $B = -3$ and $A = 2$. Substituting these values above, we get $C = 1$ and $D = -3$. Hence,

$$\frac{18+11x-x^2}{(x-1)(x+1)(x^2+3x+3)} = \frac{2}{x-1} - \frac{3}{x+1} + \frac{x-3}{x^2+3x+3}.$$

The required integral is therefore

$$\begin{aligned} \int \frac{2}{x-1} dx - \int \frac{3}{x+1} dx + \int \frac{x-3}{x^2+3x+3} dx \\ = 2 \log |x-1| - 3 \log |x+1| + \frac{1}{2} \log |x^2+3x+3| \\ - \frac{9}{\sqrt{3}} \tan^{-1} \frac{2x+3}{\sqrt{3}} + K. \end{aligned}$$

When a quadratic factor of the denominator can be broken into two real linear factors it is unnecessary to keep it in quadratic form. It may be advisable to do so, however, if the linear factors have irrational coefficients, since these coefficients make the work of calculation of the numerators more cumbersome. If it is to be kept in quadratic form, we shall need a formula for integrating fractions having this kind of quadratic denominator. The inverse tangent formula will not apply, since in this case completion of the square by adding a number to the quadratic and linear terms and subtracting the same number from the constant will lead to the difference of two squares instead of their sum. The needed formula is found by the application of the method of partial fractions in a simple case. This is done in Example 3.

EXAMPLE 3. Evaluate $\int \frac{1}{x^2 - a^2} dx$.

SOLUTION. We write

$$\frac{1}{x^2 - a^2} = \frac{1}{(x-a)(x+a)} = \frac{A}{x-a} + \frac{B}{x+a}.$$

By the method used in Example 1, we find that $A = \frac{1}{2a}$, $B = -\frac{1}{2a}$.

Hence,

$$\begin{aligned} \int \frac{1}{x^2 - a^2} dx &= \frac{1}{2a} \int \frac{1}{x-a} dx - \frac{1}{2a} \int \frac{1}{x+a} dx \\ &= \frac{1}{2a} \log |x-a| - \frac{1}{2a} \log |x+a| + C, \end{aligned}$$

$$\int \frac{1}{x^2 - a^2} dx = \frac{1}{2a} \log \left| \frac{x-a}{x+a} \right| + C.$$

This can also be written

$$\int \frac{1}{a^2 - x^2} dx = \frac{1}{2a} \log \left| \frac{x+a}{x-a} \right| + C. \quad (I-22)$$

The absolute value sign can be omitted in formula I-22 by substituting the two formulas

$$\int \frac{1}{a^2 - x^2} dx = \frac{1}{2a} \log \frac{x+a}{x-a} + C, \quad \text{if } |x| > |a|, \quad (I-22a)$$

and

$$\int \frac{1}{a^2 - x^2} dx = \frac{1}{2a} \log \frac{a+x}{a-x} + C, \quad \text{if } |x| < |a|. \quad (I-22b)$$

The distinction between these formulas is especially important when dealing with definite integrals. If the range over which the integral extends is such that $|x|$ is always greater than $|a|$, we choose (I-22a); if $|x|$ is always less than $|a|$, we choose (I-22b). If the range includes either $x = a$ or $x = -a$, the integral does not exist (see Chapter 17).

EXAMPLE 4. Evaluate $\int \frac{2x^3 + 12x^2 + 19x + 5}{(x+2)^2(x^2+3x+1)} dx$.

SOLUTION. The quadratic factor $x^2 + 3x + 1$ can be written as the product of the real linear factors $\left(x + \frac{3-\sqrt{5}}{2}\right)\left(x + \frac{3+\sqrt{5}}{2}\right)$, and we can treat this integral as that of a fraction with only linear factors in the denominator. It will be simpler, however, to leave it in quadratic form, and write

$$\frac{2x^3 + 12x^2 + 19x + 5}{(x+2)^2(x^2+3x+1)} = \frac{A}{x+2} + \frac{B}{(x+2)^2} + \frac{Cx+D}{x^2+3x+1}.$$

Proceeding as in Example 2, we have

$$\begin{aligned} A(x+2)(x^2+3x+1) + B(x^2+3x+1) + (Cx+D)(x+2)^2 \\ \equiv 2x^3 + 12x^2 + 19x + 5. \end{aligned}$$

Multiplying and collecting terms,

$$\begin{aligned} (A+C)x^3 + (5A+B+4C+D)x^2 + (7A+3B+4C+4D)x + (2A+B+4D) \\ \equiv 2x^3 + 12x^2 + 19x + 5. \end{aligned}$$

Equating coefficients as in Example 2,

$$(1) \quad A + C = 2,$$

$$(2) \quad 5A + B + 4C + D = 12,$$

$$(3) \quad 7A + 3B + 4C + 4D = 19,$$

$$(4) \quad 2A + B + 4D = 5.$$

Subtracting Eq. 4 from Eq. 2,

$$(5) \quad 3A + 4C - 3D = 7.$$

Multiplying Eq. 4 by 3 and subtracting from Eq. 3,

$$(6) \quad A + 4C - 8D = 4.$$

Subtracting Eq. 6 from Eq. 5,

$$2A + 5D = 3.$$

Multiplying Eq. 1 by 4 and subtracting from Eq. 5,

$$-A - 3D = -1.$$

Solving with

$$2A + 5D = 3,$$

we obtain $A = 4$, $D = -1$. Substituting in earlier equations, $C = -2$, $B = 1$.

The given integral is now equal to

$$\begin{aligned} & \int \frac{4}{x+2} dx + \int \frac{1}{(x+2)^2} dx - \int \frac{2x+1}{x^2+3x+1} dx \\ &= 4 \int \frac{1}{x+2} dx + \int \frac{1}{(x+2)^2} dx - \int \frac{2x+3}{x^2+3x+1} dx + 2 \int \frac{1}{x^2+3x+1} dx \\ &= 4 \log |x+2| - \frac{1}{x+2} - \log |x^2+3x+1| + 2 \int \frac{1}{x^2+3x+1} dx. \end{aligned}$$

Now,

$$\int \frac{1}{x^2+3x+1} dx = \int \frac{1}{\left(x+\frac{3}{2}\right)^2 - \left(\frac{\sqrt{5}}{2}\right)^2} dx = \int \frac{1}{u^2 - a^2} du,$$

where $u = x + 3/2$, $a = \sqrt{5}/2$. Hence, by formula I-22,

$$\int \frac{1}{x^2+3x+1} dx = \frac{1}{\sqrt{5}} \log \left| \frac{x+3/2 - \sqrt{5}/2}{x+3/2 + \sqrt{5}/2} \right| = \frac{1}{\sqrt{5}} \log \left| \frac{2x+3 - \sqrt{5}}{2x+3 + \sqrt{5}} \right|.$$

The complete integral is

$$4 \log |x+2| - \frac{1}{x+2} - \log |x^2+3x+1| + \frac{2}{\sqrt{5}} \log \left| \frac{2x+3 - \sqrt{5}}{2x+3 + \sqrt{5}} \right| + K.$$

EXERCISES

Evaluate the following integrals.

$$1. \int \frac{3x+2}{x^2+x} dx = \log(x^2|x+1|) + C.$$

$$2. \int \frac{3x dx}{(x-1)(x-2)}.$$

$$3. \int \frac{5dx}{(x+2)(2x-1)} = \log \left| \frac{2x-1}{x+2} \right| + C. \quad 4. \int \frac{(x-5)dx}{x^2-1}.$$

$$5. \int \frac{3x^2 - x + 2}{x(x^2 + 1)} dx = \log x^2 \sqrt{x^2 + 1} - \tan^{-1} x + C.$$

$$6. \int \frac{5x-6}{x(x-1)(x-2)} dx.$$

$$7. \int \frac{4x^2 - 7x + 5}{x(x-1)^2} dx = \log \left| \frac{x^5}{x-1} \right| - \frac{2}{x-1} + C.$$

$$8. \int \frac{3x^2 + x + 3}{x(x^2 + x + 1)} dx.$$

$$9. \int \frac{2(x+1)dx}{x^2(x^2+2x+2)} = -\frac{1}{x} - \tan^{-1}(x+1) + C.$$

$$10. \int \frac{10x^2 + 21x + 38}{(x^2+4)(x^2+x-6)} dx.$$

$$11. \int \frac{7x^2 + 45x + 80}{x^3 + 8x^2 + 16x} dx = 5 \log |x| + 2 \log |x+4| + \frac{3}{x+4} + C.$$

$$12. \int \frac{1-2x-x^2}{x^2(x+1)} dx.$$

$$13. \int \frac{(5x^3 + 9x + 6)dx}{(x+1)^2(x^2+3)} = 5 \log |x+1| + \frac{2}{x+1} - \sqrt{3} \tan^{-1} \frac{x}{\sqrt{3}} + C.$$

$$14. \int \frac{(2x^4 - 5x^3 + x^2 + x - 5)dx}{(x-1)^2(x^2+1)}.$$

$$15. \int \frac{dx}{x^3+1} = \frac{1}{3} \log |x+1| - \frac{1}{6} \log |x^2-x+1| + \frac{1}{\sqrt{3}} \tan^{-1} \frac{2x-1}{\sqrt{3}} + C.$$

$$16. \int_0^1 \frac{dx}{4-x^2}.$$

$$17. \int_4^{10} \frac{x dx}{(x-1)(x-3)} = \log \frac{7\sqrt{21}}{3}.$$

$$18. \int_5^9 \frac{3 dx}{x^2-9}.$$

$$19. \int_2^5 \frac{4 dx}{1-x^2} = -\log 4.$$

$$20. \int_{-1}^1 \frac{dx}{x^2-4}.$$

7. Integration by Parts. One of the most useful rules of differentiation is that by which we differentiate products. This rule has not been inverted to produce an integration formula. The reason for this is simply that such a formula would require the presence together of two closely related terms, which it is unreasonable to expect, except as the immediate result of differentiating a product. For example,

$$\frac{d}{dx}(x \sin x) = x \cos x + \sin x,$$

and, of course,

$$\int (x \cos x + \sin x) dx = x \sin x + C.$$

This result has little use as a formula, since the integrand is a special combination of terms unlikely to be encountered. Nevertheless, this example illustrates a procedure that has wide application.

The two terms needed to make the above integration possible are $x \cos x$ and $\sin x$. The first of these alone falls under none of the integration formulas listed in Art. 3. The second fits formula I-3. From the equation

$$\int (x \cos x + \sin x) dx = x \sin x + C$$

we have

$$\int x \cos x dx + \int \sin x dx = x \sin x + C.$$

Transposing,

$$\begin{aligned} \int x \cos x dx &= x \sin x - \int \sin x dx + C \\ &= x \sin x - \cos x + C. \end{aligned}$$

Hence, since one of the terms arising from product differentiation can be integrated by itself, we are able to integrate the other.

In general terms, the above method can be written as follows. If u and v are two functions of x , $u(x)$, and $v(x)$,

$$\frac{d}{dx}(uv) = u \frac{dv}{dx} + v \frac{du}{dx}$$

Hence,

$$\int \left(u \frac{dv}{dx} + v \frac{du}{dx} \right) dx = uv + C.$$

That is,

$$\int \left(u \frac{dv}{dx} \right) dx + \int \left(v \frac{du}{dx} \right) dx = uv + C.$$

Transposing,

$$\int \left(u \frac{dv}{dx} \right) dx = uv - \int \left(v \frac{du}{dx} \right) dx + C. \quad (7.1)$$

In differential form,

$$\int u \, dv = uv - \int v \, du + C. \quad (7.2)$$

The separation of an integral into parts so as to apply the above method is illustrated in the following examples.

EXAMPLE 1. $\int x e^x \, dx$.

SOLUTION. We identify x with u and e^x with dv/dx . Then $v = e^x$ and $du/dx = 1$. Hence, by (7.1),

$$\begin{array}{ccccccc} \int x & e^x & dx & = & x e^x & - & \int e^x \cdot 1 \, dx + C. \\ \uparrow & \uparrow & & & \uparrow & \uparrow & \uparrow \\ u & dv/dx & & & u & v & v \, du/dx \end{array}$$

When obtaining v from dv/dx , it is not necessary to include a constant of integration, since we only require *some* function whose derivative is dv/dx . The student can show that inclusion of such a constant has no effect on the final result in this example.

EXAMPLE 2. $\int x \log x \, dx$.

SOLUTION. We wish to think of this as having the form $\int u \frac{dv}{dx} \, dx$, in order to fit it to formula 7.1. It will not do to put $dv/dx = \log x$, since we shall then have no means of finding v . Instead, we let

$$\log x = u, \quad \text{whence} \quad du/dx = 1/x,$$

and

$$x = dv/dx, \quad v = x^2/2.$$

Thus,

$$\begin{array}{ccccccc} \int x & \log x & dx & = & x^2/2 \cdot \log x & - & \int x^2/2 \cdot 1/x \, dx + C \\ \uparrow & \uparrow & & & \uparrow & \uparrow & \uparrow \\ dv/dx & u & & & v & u & v \, du/dx \end{array}$$

$$= (x^2/2) \log x - (x^2/4) + C.$$

EXAMPLE 3. $\int x^2 \sin x \, dx$.

SOLUTION. In formula 7.1, the quantity that we select for u is differentiated, and the other factor is dv/dx and is integrated to give v . In the present example, differentiation of x^2 reduces the degree whereas integration increases it. Either of these operations changes $\sin x$ into $\cos x$. It seems most promising, therefore, to let $u = x^2$ and $dv/dx = \sin x$, since thereby we are left with an integration to be performed upon a simpler function than the one we start with.

$$\begin{array}{ccccccc} \int x^2 & \sin x & dx & = & x^2 (-\cos x) & - & \int 2x (-\cos x) \, dx + C \\ \uparrow & \uparrow & & & \uparrow & \uparrow & \uparrow \\ u & dv/dx & & & u & v & du/dx \end{array}$$

$$= -x^2 \cos x + 2 \int x \cos x \, dx + C. \quad (7.3)$$

We can now apply the same process to $\int x \cos x \, dx$

$$\begin{aligned} \int x \cos x \, dx &= x \sin x - \int 1 \cdot \sin x \, dx + C' \\ \uparrow \quad \uparrow &\quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ u \, dv/dx &\quad u \quad v \quad du/dx \quad v \\ &= x \sin x + \cos x + C'. \end{aligned}$$

Inserting this result in (7.3), we have

$$\int x^3 \sin x \, dx = -x^2 \cos x + 2x \sin x + 2 \cos x + K,$$

where $K = C - C'$. It is clearly unnecessary to insert two arbitrary constants. It is sufficient to annex one constant when the integration is completed.

This is an example of repeated integration by parts. The student can easily show that the process would not yield the result if we had chosen u to be $\sin x$ and dv/dx to be x^2 . The rule here followed was to select for u a factor that is simplified upon differentiation. At the same time, the other factor, dv/dx , must be a function that can be integrated. Experience rather than any single rule, however, is the best guide in determining the breakup into factors.

EXAMPLE 4. $\int \tan^{-1} x \, dx$.

SOLUTION. We can think of the integrand as the product $1 \cdot \tan^{-1} x$. If we now put $u = \tan^{-1} x$, du/dx is the algebraic function $1/(1+x^2)$. At the same time, $dv/dx = 1$, and $v = x$. The second term in the integration is $\int \frac{x}{1+x^2} \, dx$, which is easily integrated. Thus,

$$\begin{aligned} \int 1 \cdot \tan^{-1} x \, dx &= x \cdot \tan^{-1} x - \int x \cdot \frac{1}{1+x^2} \, dx \\ \uparrow \quad \uparrow &\quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ dv/dx \quad u &\quad v \quad u \quad v \quad du/dx \\ &= x \tan^{-1} x - \frac{1}{2} \log(1+x^2) + C. \end{aligned}$$

EXAMPLE 5. $\int e^{ax} \sin bx \, dx$.

SOLUTION. Here we have no guide for determining u and dv/dx . In fact, we can put $u = e^{ax}$ or $u = \sin bx$ with equal success. Letting $u = e^{ax}$, we have

$$\begin{aligned} \int e^{ax} \sin bx \, dx &= e^{ax} \left(-\frac{1}{b} \cos bx \right) - \int a e^{ax} \left(-\frac{1}{b} \cos bx \right) dx \\ \uparrow \quad \uparrow &\quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ u \, dv/dx &\quad u \quad v \quad du/dx \quad v \\ &= -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b} \int e^{ax} \cos bx \, dx. \end{aligned} \quad (7.4)$$

This result is discouraging, since we still have to evaluate an integral much like the one with which we started. If we apply integration by parts again, however, we shall find that the problem can be solved in an unexpected way. We have

$$\begin{aligned} \int e^{ax} \cos bx \, dx &= \overset{\substack{\uparrow \\ u}}{e^{ax}} \overset{\substack{\uparrow \\ dv/dx}}{\cos bx} = \overset{\substack{\uparrow \\ u}}{e^{ax}} \overset{\substack{\uparrow \\ v}}{\frac{1}{b} \sin bx} - \int \overset{\substack{\uparrow \\ du/dx}}{ae^{ax}} \overset{\substack{\uparrow \\ v}}{\frac{1}{b} \sin bx} \, dx \\ &= \frac{1}{b} e^{ax} \sin bx - \frac{a}{b} \int e^{ax} \sin bx \, dx. \end{aligned}$$

Apparently, we are now back where we started, with the same integral to be evaluated. The insertion of the last result in (7.4), however, produces an equation from which the required integral can be found.

$$\begin{aligned} \int e^{ax} \sin bx \, dx &= -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b} \left[\frac{1}{b} e^{ax} \sin bx - \frac{a}{b} \int e^{ax} \sin bx \, dx \right] \\ &= -\frac{1}{b} e^{ax} \cos bx + \frac{a}{b^2} e^{ax} \sin bx - \frac{a^2}{b^2} \int e^{ax} \sin bx \, dx. \end{aligned}$$

Transposing, and combining the two integrals,

$$\frac{a^2 + b^2}{b^2} \int e^{ax} \sin bx \, dx = \frac{e^{ax}}{b^2} (a \sin bx - b \cos bx).$$

Hence,

$$\int e^{ax} \sin bx \, dx = \frac{e^{ax}(a \sin bx - b \cos bx)}{a^2 + b^2} + C.$$

EXERCISES

Calculate the following integrals.

1. $\int x \sec^2 x \, dx = x \tan x - \log |\sec x| + C.$
2. $\int x \cos 2x \, dx.$
3. $\int \log |x| \, dx = x \log |x| - x + C.$
4. $\int x e^{3x} \, dx.$
5. $\int x^3 e^x \, dx = e^x(x^3 - 3x^2 + 6x - 6) + C.$
6. $\int x a^x \, dx.$
7. $\int t \tan^{-1} t \, dt = \frac{1}{2}(t^2 + 1) \tan^{-1} t - \frac{1}{2}t + C.$
8. $\int \cos^{-1} t \, dt.$
9. $\int y^2 \cos y \, dy = y^2 \sin y + 2y \cos y - 2 \sin y + C.$
10. $\int x^2 e^{-x} \, dx.$
11. $\int x^2 \cot^{-1} x \, dx = \frac{1}{3}x^3 \cot^{-1} x + \frac{1}{6}x^2 - \frac{1}{6} \log(x^2 + 1) + C.$

12. $\int y^{10} \log y \, dy.$
13. $\int x/\sqrt{1+x} \, dx = 2x\sqrt{1+x} - \frac{4}{3}(\sqrt{1+x})^3 + C.$
14. $\int (\log \frac{1}{t})/(1+t)^2 \, dt.$
15. $\int \theta \csc 2\theta \cot 2\theta \, d\theta = -\frac{1}{2}\theta \csc 2\theta + \frac{1}{4} \log |\csc 2\theta - \cot 2\theta| + C.$
16. $\int (e^x - x)^2 \, dx.$
17. $\int x(e^x - e^{-x}) \, dx = x(e^x + e^{-x}) - (e^x - e^{-x}) + C.$
18. $\int e^{ax} \cos bx \, dx.$
19. $\int (\log x)/\sqrt{x} \, dx = 2\sqrt{x}(\log x - 2) + C.$
20. $\int \sqrt{x} \log x \, dx.$
21. $\int \sec^3 x \, dx = \frac{1}{2}[\sec x \tan x + \log |\sec x + \tan x|] + C.$

[Hint: Write $\sec^3 x = \sec x \cdot \sec^2 x$.

Then

$$\begin{aligned} \int \sec^3 x \, dx &= \sec x \tan x - \int \sec x \tan^2 x \, dx \\ &= \sec x \tan x + \int \sec x \, dx - \int \sec^3 x \, dx. \end{aligned}$$

Now transpose $\int \sec^3 x \, dx$]

22. $\int \csc^3 \theta \, d\theta.$
23. $\int xe^x/(1+x)^2 \, dx = e^x/(1+x) + C.$ [Hint: Let $u = xe^x$.]
24. $\int \sin(\log |x|) \, dx = \frac{1}{2}x[\sin(\log |x|) - \cos(\log |x|)] + C.$
25. $\int_0^1 \sin^{-1} t \, dt = \pi/2 - 1.$
26. $\int_0^1 e^{-.01t} \sin \pi t \, dt = \frac{\pi}{\pi^2 + .01} (e^{-.1} + 1).$
27. $\int_0^{\pi/4} \sec \theta \tan^2 \theta \, d\theta = \frac{1}{2}(\sqrt{2} - \log(\sqrt{2} + 1)).$
28. $\int_1^3 x^5 \log |x| \, dx.$

8. Integration of Trigonometric Functions. Formulas I-2, I-3, and I-18 to I-21 enable us to integrate all the six fundamental trigonometric functions. In addition, formulas I-4 to I-7 dispose of some products. In the present article we shall first consider other powers and products. We shall study these by means of examples. The generalizations will be pointed out, and are readily grasped from the illustrative examples.

EXAMPLE 1. $\int \sin^3 x \cos^2 x \, dx.$

SOLUTION. We write

$$\begin{aligned} \sin^3 x \cos^2 x &= \sin^2 x \cos^2 x \sin x \\ &= (1 - \cos^2 x) \cos^2 x \sin x \\ &= \cos^2 x \sin x - \cos^4 x \sin x. \end{aligned}$$

Hence,

$$\begin{aligned}\int \sin^3 x \cos^2 x \, dx &= \int \cos^2 x \sin x \, dx - \int \cos^4 x \sin x \, dx \\ &= -\frac{1}{3} \cos^3 x + \frac{1}{5} \cos^5 x + C,\end{aligned}$$

by the general power formula.

GENERALIZATION. This method can be used with all functions of forms $\sin^n x$, $\cos^n x$, $\sin^n x \cos^m x$, and $\sin^m x \cos^n x$, provided that n is odd.

EXAMPLE 2. $\int \sin^4 x \, dx$.

SOLUTION. We write

$$\begin{aligned}\sin^4 x &= (\sin^2 x)^2 = \left(\frac{1}{2} - \frac{1}{2} \cos 2x\right)^2 \\ &= \frac{1}{4}(1 - 2 \cos 2x + \cos^2 2x) \\ &= \frac{1}{4} - \frac{1}{2} \cos 2x + \frac{1}{4}\left(\frac{1}{2} + \frac{1}{2} \cos 4x\right) \\ &= \frac{3}{8} - \frac{1}{2} \cos 2x + \frac{1}{8} \cos 4x.\end{aligned}$$

Hence,

$$\begin{aligned}\int \sin^4 x \, dx &= \int \frac{3}{8} \, dx - \int \frac{1}{2} \cos 2x \, dx + \int \frac{1}{8} \cos 4x \, dx \\ &= \frac{3}{8}x - \frac{1}{4} \sin 2x + \frac{1}{32} \sin 4x + C.\end{aligned}$$

GENERALIZATION. This method can be used with all even powers of $\sin x$ and $\cos x$ and products of such powers. The method of Example 1 will not work in these cases.

EXAMPLE 3. $\int \tan^3 x \sec^4 x \, dx$.

SOLUTION. We write $\tan^3 x \sec^4 x = \tan^3 x(1 + \tan^2 x) \sec^2 x$, reserving $\sec^2 x$ in this transformation, since it is the derivative of $\tan x$. Hence,

$$\begin{aligned}\int \tan^3 x \sec^4 x \, dx &= \int \tan^3 x \sec^2 x \, dx + \int \tan^5 x \sec^2 x \, dx \\ &= \frac{1}{4} \tan^4 x + \frac{1}{6} \tan^6 x + C,\end{aligned}$$

by the general power formula.

GENERALIZATION. This method can be used with all products of the form $\tan^n x \sec^m x$, provided that m is even, including $n = 0$; that is, even powers of $\sec x$. A method similar in all respects can be applied to functions of the form $\cot^n x \csc^m x$, if m is even.

EXAMPLE 4. $\int \tan^3 x \sec^3 x \, dx$.

SOLUTION. We write

$$\begin{aligned}\tan^3 x \sec^3 x &= \tan^2 x \sec^2 x (\tan x \sec x) \\ &= (\sec^2 x - 1) \sec^2 x (\tan x \sec x)\end{aligned}$$

separating the factor $\tan x \sec x$ from the rest of the function, since it is the derivative of $\sec x$. Hence,

$$\begin{aligned}\int \tan^3 x \sec^3 x \, dx &= \int \sec^4 x (\tan x \sec x) \, dx - \int \sec^2 x (\tan x \sec x) \, dx \\ &= \frac{1}{5} \sec^5 x - \frac{1}{3} \sec^3 x + C,\end{aligned}$$

by the general power formula.

GENERALIZATION. This method can be used with all products of the form $\tan^n x \sec^m x$, provided that n is odd, including $m = 0$; that is, odd powers of $\tan x$. A similar method can be applied to functions of the form $\csc^n x \sec^m x$, if n is odd.

EXAMPLE 5. $\int \sec^5 x \, dx$.

SOLUTION. Writing $\int \sec^5 x \, dx = \int \sec^3 x \sec^2 x \, dx$, we can integrate by parts as follows.

$$\begin{aligned} \int \sec^3 x \sec^2 x \, dx &= \sec^3 x \tan x - \int (3 \sec^2 x \sec x \tan x) \tan x \, dx \\ \begin{array}{ccccccc} \uparrow & \uparrow & & \uparrow & & \uparrow & \\ u & dv/dx & & u & & dv/dx & \\ & & & v & & & \end{array} \\ &= \sec^3 x \tan x - 3 \int \sec^3 x \tan^2 x \, dx \\ &= \sec^3 x \tan x - 3 \int \sec^3 x (\sec^2 x - 1) \, dx. \end{aligned}$$

Hence,

$$\int \sec^5 x \, dx = \sec^3 x \tan x - 3 \int \sec^5 x \, dx + 3 \int \sec^3 x \, dx.$$

Transposing $3 \int \sec^5 x \, dx$,

$$4 \int \sec^5 x \, dx = \sec^3 x \tan x + \frac{3}{2} \sec x \tan x + \frac{3}{2} \log (\sec x + \tan x),$$

by the result of Exercise 21, Art. 7. Therefore, dividing by 4,

$$\int \sec^5 x \, dx = \frac{1}{4} \sec^3 x \tan x + \frac{3}{8} \sec x \tan x + \frac{3}{8} \log (\sec x + \tan x) + C.$$

GENERALIZATION. This method can be applied to any product of the form $\sec^n x \tan^m x$, or $\csc^n x \csc^m x$, if n is odd and m is even, including $m = 0$; i.e., odd powers of $\sec x$ or $\csc x$, as in the example.

It will be seen that the methods of Examples 1 and 2 cover all products of positive integral powers of $\sin x$ and $\cos x$, whereas the methods of Examples 3, 4, and 5 cover all products of positive integral powers of $\tan x$ and $\sec x$, and of $\csc x$ and $\csc x$.

EXAMPLE 6. $\int \sin 2x \cos 3x \, dx$.

SOLUTION. The product $\sin 2x \cos 3x$ can be written as an algebraic sum of single trigonometric terms by means of formula 8.1 below. Thus,

$$\sin 2x \cos 3x = \frac{1}{2} (\sin 5x - \sin x).$$

Hence,

$$\begin{aligned} \int \sin 2x \cos 3x \, dx &= \frac{1}{2} \int \sin 5x \, dx - \frac{1}{2} \int \sin x \, dx \\ &= -\frac{1}{10} \cos 5x + \frac{1}{2} \cos x + C. \end{aligned}$$

GENERALIZATION. This method can be used with all products of the type $\sin Mx \cos Nx$, $\sin Mx \sin Nx$, and $\cos Mx \cos Nx$. In fact, by making use of the trigonometric formulas

$$\sin A \cos B = \frac{1}{2} [\sin(A + B) + \sin(A - B)], \quad (8.1)$$

$$\sin A \sin B = -\frac{1}{2} [\cos(A + B) - \cos(A - B)], \quad (8.2)$$

$$\cos A \cos B = \frac{1}{2} [\cos(A + B) + \cos(A - B)], \quad (8.3)$$

we have

$$\int \sin Mx \cos Nx \, dx$$

$$= -\frac{1}{2(M+N)} \cos(M+N)x - \frac{1}{2(M-N)} \cos(M-N)x + C,$$

$$\int \sin Mx \sin Nx \, dx$$

$$= -\frac{1}{2(M+N)} \sin(M+N)x + \frac{1}{2(M-N)} \sin(M-N)x + C,$$

$$\int \cos Mx \cos Nx \, dx$$

$$= \frac{1}{2(M+N)} \sin(M+N)x + \frac{1}{2(M-N)} \sin(M-N)x + C.$$

EXERCISES

Calculate the following integrals.

1. $\int \cos^3 x \, dx = \sin x - \frac{1}{3} \sin^3 x + C.$
2. $\int \sin^2 x \cos^2 x \, dx.$
3. $\int \sin^5 x \, dx = -\cos x + \frac{2}{3} \cos^3 x - \frac{1}{5} \cos^5 x + C.$
4. $\int \cos 2x \sin^3 2x \, dx.$
5. $\int \sin^2 \frac{x}{2} \cos^5 \frac{x}{2} \, dx = \frac{2}{3} \sin^3 \frac{x}{2} - \frac{4}{5} \sin^5 \frac{x}{2} + \frac{2}{7} \sin^7 \frac{x}{2} + C.$
6. $\int \frac{\cos x}{\sin^3 x} \, dx.$
7. $\int \frac{\sin x}{\sqrt{\cos x}} \, dx = 2\sqrt{\cos x} + C.$
8. $\int \frac{\cos t}{\sqrt{1 - \sin t}} \, dt.$
9. $\int \frac{1 - \cos t}{\sin^2 t} \, dt = -\cot t + \csc t + C.$
10. $\int \frac{1 - \cos y}{1 + \cos y} \, dy.$
11. $\int \sin^2 \theta \cos^2 \theta \, d\theta = \frac{1}{8} \theta - \frac{1}{32} \sin 4\theta + C.$
12. $\int \cos^4 x \sin^2 x \, dx.$
13. $\int \tan^3 x \, dx = \frac{1}{2} \tan^2 x - \log |\sec x| + C.$
14. $\int \cot^2 t \csc^2 t \, dt.$
15. $\int \frac{\cos^3 x}{\sin^4 x} \, dx = -\frac{1}{3} \csc^3 x + \csc x + C.$

$$16. \int \csc^4 2x \, dx.$$

$$17. \int \sec^4 \frac{t}{2} \, dt = \frac{2}{3} \tan^3 \frac{t}{2} + 2 \tan \frac{t}{2} + C.$$

$$18. \int \frac{\sin x - 1}{\cos^3 x} \, dx.$$

$$19. \int \tan^3 x \sqrt{\sec x} \, dx = 2\sqrt{\sec x} \left(\frac{1}{3} \sec^2 x - 1 \right) + C.$$

[Hint: Write $\sqrt{\sec x} = \sec x / \sqrt{\sec x}$.]

$$20. \int \frac{\sec^4 t}{\sqrt{\tan t}} \, dt.$$

$$21. \int \sin 2x \cos x \, dx = -\frac{1}{8}(\cos 3x + 3 \cos x) + C.$$

$$22. \int \cos 5x \sin 3x \, dx.$$

$$23. \int (\sqrt{\sin \theta} - \cos \theta)^2 \, d\theta = \frac{1}{2}\theta - \cos \theta - \frac{4}{3} \sin \theta \sqrt{\sin \theta} + \frac{1}{4} \sin 2\theta + C.$$

$$24. \int (\cos \theta - 2 \cos 2\theta)^2 \, d\theta.$$

$$25. \int \csc x \csc^2 x \, dx = -\frac{1}{2}(\csc x \cot x + \log |\csc x - \cot x|) + C.$$

$$26. \int \csc^5 x \, dx.$$

$$27. \int_{\frac{1}{2}}^{\frac{1}{3}} \frac{\cos^2 \pi x}{\sin^4 \pi x} \, dx = \sqrt{3}/\pi.$$

$$28. \int_{\pi/6}^{\pi/4} \sec^2 \theta \csc^4 \theta \, d\theta.$$

$$29. \int_0^{\pi} \sin^2 3x \cos^2 x \, dx = \pi/4.$$

$$30. \int_0^{\pi/3} \frac{\tan^5 x}{\sec^3 x} \, dx.$$

9. Trigonometric Substitutions. The integration of many irrational expressions can be dealt with successfully by means of trigonometric substitutions. This method is illustrated in the following examples.

EXAMPLE 1. $\int \sqrt{1-x^2} \, dx$, $|x| < 1$.

SOLUTION. We make the substitution $u = \sin^{-1} x$ (principal value), whence $-\pi/2 < u < \pi/2$, and $\cos u > 0$. By formula 2.1 for change of variable, with $x = \sin u$, $\sqrt{1-x^2} = \cos u$, $dx = \cos u \, du$, we have

$$\begin{aligned} \int \sqrt{1-x^2} \, dx &= \int \cos u \cos u \, du \\ &= \int \left(\frac{1}{2} + \frac{1}{2} \cos 2u \right) du \\ &= \frac{1}{2}u + \frac{1}{4} \sin 2u \\ &= \frac{1}{2}u + \frac{1}{2} \sin u \cos u \\ &= \frac{1}{2} \sin^{-1} x + \frac{1}{2} x \sqrt{1-x^2} + C. \end{aligned}$$

EXAMPLE 2. $\int 1/\sqrt{4+x^2} dx$.

SOLUTION. A trigonometric substitution that will rationalize this expression is $u = \tan^{-1}(x/2)$, with $-\pi/2 < u < \pi/2$. In this range $\sec u$ is positive, and

$$x = 2 \tan u, \quad \sqrt{4+x^2} = 2 \sec u, \quad dx = 2 \sec^2 u du.$$

Hence,

$$\begin{aligned} \int 1/\sqrt{4+x^2} dx &= \int \sec^2 u / \sec u du = \int \sec u du \\ &= \log(\sec u + \tan u) + C \\ &= \log(x + \sqrt{4+x^2}) + C. \end{aligned}$$

Here we let C absorb the constant, $-\log 2$.

As a generalization of this result, we have

$$\int 1/\sqrt{x^2+a^2} dx = \log(x + \sqrt{x^2+a^2}) + C, \quad (I-23)$$

which is stated among the formulas for integration in Art. 3. The result,

$$\int 1/\sqrt{x^2-a^2} dx = \log(x + \sqrt{x^2-a^2}) + C, \quad (I-24)$$

can be obtained by the substitution $u = \sec^{-1}(x/a)$. See Exercise 7 below.

EXAMPLE 3. $\int \sqrt{x^2-9} dx, |x| > 3$.

SOLUTION. For $x > 3$, this integrand can be rationalized by the substitution, $u = \sec^{-1}(x/3)$, $0 < u < \pi/2$. Then

$$x = 3 \sec u, \quad \sqrt{x^2-9} = 3 \tan u, \quad dx = 3 \sec u \tan u du.$$

Hence,

$$\int \sqrt{x^2-9} dx = \int 9 \tan^2 u \sec u du.$$

This is one of those troublesome trigonometric integrals like $\int \sec^3 x dx$, which can be handled as in Example 5, Art. 8. By this method we obtain

$$\int \tan^2 u \sec u du = \frac{1}{2}[\tan u \sec u - \log |\sec u + \tan u|] + C,$$

whence

$$\begin{aligned} \int \sqrt{x^2-9} dx &= \frac{9}{2}[\sqrt{x^2-9}/3 \cdot (x/3) - \log |x/3 + \sqrt{x^2-9}/3|] + C \\ &= \frac{1}{2}[x\sqrt{x^2-9} - 9 \log |x + \sqrt{x^2-9}|] + K, \end{aligned}$$

in which $K = C + \frac{9}{2} \log 3$.

If $x < -3$, we can use the substitution $u = \sec^{-1}(x/3)$ with $\pi < u < 3\pi/2$, so that $\tan u$ is positive, as required by the positive radical $\sqrt{x^2 - 9}$. The details of the substitution are the same as before.

Our experience with these examples shows us that

$\sqrt{a^2 - x^2}$ is rationalized by putting $u = \sin^{-1}(x/a)$, or $x = a \sin u$,

$\sqrt{a^2 + x^2}$ is rationalized by putting $u = \tan^{-1}(x/a)$, or $x = a \tan u$,

$\sqrt{x^2 - a^2}$ is rationalized by putting $u = \sec^{-1}(x/a)$, or $x = a \sec u$.

Figures 1, 2, and 3 show the relations between the variables x and u in these three substitutions and may be helpful both in making the substitution and in replacing u by x in the result.

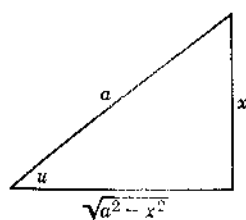


FIG. 1

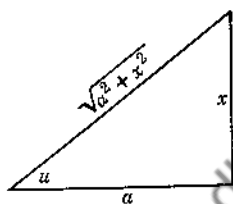


FIG. 2

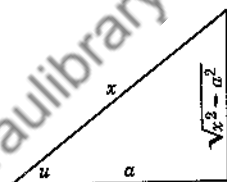


FIG. 3

We give one more example to illustrate the use of this type of substitution in the evaluation of a definite integral.

EXAMPLE 4. Evaluate $\int_0^5 \sqrt{x^2 + 25} dx$.

SOLUTION. Putting $u = \tan^{-1}(x/5)$, we have

$$x = 5 \tan u, \quad \sqrt{x^2 + 25} = 5 \sec u, \quad dx = 5 \sec^2 u du.$$

The new limits in the transformed integral are found as follows: if $x = 0$, $u = \tan^{-1} 0 = 0$, and if $x = 5$, $u = \tan^{-1} 1 = \pi/4$. Hence,

$$\int_0^5 \sqrt{x^2 + 25} dx = \int_0^{\pi/4} 25 \sec^3 u du.$$

Using the result found for this integral in Exercise 21, Art. 7, we have

$$\begin{aligned} \int_0^5 \sqrt{x^2 + 25} dx &= \frac{25}{2} \left[\tan u \sec u + \log (\sec u + \tan u) \right]_0^{\pi/4} \\ &= \frac{25}{2} [1\sqrt{2} + \log (\sqrt{2} + 1) - 0 - \log (1 + 0)] \\ &= \frac{25}{2} [\sqrt{2} + \log (\sqrt{2} + 1)]. \end{aligned}$$

EXERCISES

Evaluate the following integrals.

1. $\int \frac{dx}{(1-x^2)^{\frac{3}{2}}} = \frac{x}{\sqrt{1-x^2}} + C.$
2. $\int \frac{dx}{x\sqrt{4-x^2}}.$
3. $\int \frac{dx}{(x^2+9)^{\frac{3}{2}}} = \frac{x}{9\sqrt{x^2+9}} + C.$
4. $\int \frac{\sqrt{s^2-1}}{s} ds.$
5. $\int \frac{\sqrt{25-x^2}}{x^2} dx = -\frac{\sqrt{25-x^2}}{x} - \sin^{-1} \frac{x}{5} + C.$
6. $\int \frac{r^2 dr}{(9-r^2)^{\frac{3}{2}}}.$
7. $\int \frac{dx}{\sqrt{x^2-a^2}} = \log(x + \sqrt{x^2-a^2}) + C.$
8. $\int \frac{dx}{x^2\sqrt{x^2-5}}.$
9. $\int \frac{\sqrt{16-t^2}}{t} dt = 4 \log |(4 - \sqrt{16-t^2})/t| + \sqrt{16-t^2} + C.$
10. $\int \frac{dz}{z^2\sqrt{z^2-9}}.$
11. $\int \frac{dx}{x^2\sqrt{9-x^2}} = \frac{1}{54} [\log |(3 - \sqrt{9-x^2})/x| - 3\sqrt{9-x^2}/x^2] + C.$
12. $\int \frac{x^2 dx}{\sqrt{x^2-1}}.$
13. $\int_{\sqrt{2}}^2 \frac{dx}{x^2\sqrt{x^2-1}} = \frac{1}{2}(\sqrt{3} - \sqrt{2}).$
14. $\int_5^{10} \frac{\sqrt{x^2-25}}{x} dx.$
15. $\int_2^{2\sqrt{3}} \frac{dx}{x\sqrt{16-x^2}} = \frac{1}{4} \log \frac{2+\sqrt{3}}{\sqrt{3}}.$
16. $\int_1^2 \frac{dx}{x(x^2+4)^{\frac{3}{2}}}.$

10. Other Substitutions. The variety of substitutions that can be employed for the reduction of integrals is endless. In a first course we can become acquainted with only a few of the most useful ones. The student will find others in books that devote more attention to the development of techniques. He may also find helpful substitutions for himself, by trial. In the present article we shall take up some examples further illustrating the method.

Although a substitution may lead to a result that is formally correct, it may at the same time limit the range within which the result is valid. For example, the substitution $u = \sqrt{x}$ is permissible only for positive values of x ; and $u = \sin^{-1} x$ is valid only in the range $-1 \leq x \leq 1$. Unless care is exercised, a substitution may lead to an incorrect result in a definite integral; e.g., if $u = (x+1)/x$, and $x = 0$ lies in the interval of integration, as illustrated in Example 3.

EXAMPLE 1.
$$\int \frac{x \, dx}{\sqrt{x+1} + \sqrt[3]{x+1}}$$

SOLUTION. Since we have the square root and the cube root of the same expression, this integrand will be rationalized by setting

$$x+1 = u^6, \quad x = u^6 - 1, \quad dx = 6u^5 \, du, \quad \sqrt{x+1} = u^3, \quad \sqrt[3]{x+1} = u^2.$$

The integral then becomes

$$\begin{aligned} \int \frac{(u^6 - 1)6u^5 \, du}{u^3 + u^2} &= \int \frac{6u^3(u^6 - 1) \, du}{u + 1} \\ &= \int 6u^3(u^5 + u^4 + u^3 + u^2 + u + 1) \, du \\ &= 6u^9/9 + 6u^8/8 + 6u^7/7 + 6u^6/6 + 6u^5/5 + 6u^4/4 + C \\ &= 6u^4(u^5/9 + u^4/8 + u^3/7 + u^2/6 + u/5 + 1/4) + C. \end{aligned}$$

Since $u = \sqrt[6]{x+1}$, this is equal to

$$\begin{aligned} 6\sqrt[6]{(x+1)^4} & \left[\sqrt[6]{(x+1)^5}/9 + \sqrt[6]{(x+1)^4}/8 + \sqrt[6]{(x+1)^3}/7 \right. \\ & \left. + \sqrt[6]{(x+1)^2}/6 + \sqrt[6]{x+1}/5 + 1/4 \right] + C. \end{aligned}$$

This method will work with integrands that are irrational only because of the presence of one or more fractional powers of the same expression, like $(x+1)$ in the example.

EXAMPLE 2.
$$\int \frac{\sqrt[3]{x+x^3}}{x^4} \, dx.$$

SOLUTION. The substitution, $x = 1/u$, will have the effect of placing the first power of u in the numerator. Since, after the substitution, the

radical will be $\sqrt[3]{u^2 + 1}$, the resulting integral can be readily calculated. Hence, we put $x = 1/u$, $dx = -du/u^2$, and the integral becomes

$$\begin{aligned}\int \frac{\sqrt[3]{x + x^3}}{x^4} dx &= \int u^4 \sqrt[3]{1/u + 1/u^3} (-1/u^2) du \\ &= -\int \sqrt[3]{u^2 + 1} u du \\ &= -\frac{1}{2} \cdot \frac{2}{3} (u^2 + 1)^{3/2} + C \\ &= -\frac{2}{3} (1/x^2 + 1)^{3/2} + C.\end{aligned}$$

The substitution, $x = 1/u$, is useful for transforming many irrational integrands. No general rules for its use can be laid down. It usually has the effect, as in this example, of transforming one irrational integrand into another that is also irrational, but the second can often be integrated more easily than the first.

EXAMPLE 3. $\int_{-1}^{1/7} \frac{dx}{x \sqrt[3]{x(x+1)^2}}$

SOLUTION. Letting $x = 1/(u-1)$, we have

$$x+1 = u/(u-1), \quad dx = -du/(u-1)^2, \quad \text{and } u = (x+1)/x.$$

This substitution leads formally to the following.

$$\begin{aligned}\int_{-1}^{1/7} \frac{dx}{x \sqrt[3]{x(x+1)^2}} &= \int_0^8 \frac{-du/(u-1)^2}{1/(u-1) \cdot \sqrt[3]{1/(u-1) \cdot u^2/(u-1)^2}} \\ &= \int_0^8 -\frac{du}{u^2} = \left[-3u^3 \right]_0^8 = -6.\end{aligned}$$

This result is incorrect since the substitution $u = (x+1)/x$ is not valid for $x = 0$, which lies in the interval of integration. The discussion in Chapter 17 will show that the given definite integral has no meaning. In an interval excluding $x = 0$, however, the substitution is valid. The student can readily show that the primitive function obtained, viz., $-3[(x+1)/x]^3$, is formally correct.

EXAMPLE 4. $\int \frac{1}{2 \cos x + \sin x + 3} dx.$

SOLUTION. This integrand is transformed into a rational algebraic fraction by any substitution that replaces all of the quantities $\cos x$, $\sin x$, and dx by rational algebraic fractions. Since

$$(1 - u^2)^2 + (2u)^2 = (1 + u^2)^2, \quad \text{and} \quad \left(\frac{1 - u^2}{1 + u^2} \right)^2 + \left(\frac{2u}{1 + u^2} \right)^2 = 1,$$

we see from Fig. 4 that it is possible to make

$$\sin x = \frac{2u}{1+u^2} \quad \text{and} \quad \cos x = \frac{1-u^2}{1+u^2} \quad (10.1)$$

Differentiating the first of these relations,

$$\begin{aligned} \cos x \, dx &= \frac{(1+u^2) \cdot 2 - 2u \cdot 2u}{(1+u^2)^2} du = 2 \cdot \frac{1-u^2}{1+u^2} \cdot \frac{du}{1+u^2} \\ &= 2 \cos x \frac{du}{1+u^2} \end{aligned}$$

Hence,

$$dx = \frac{2du}{1+u^2}.$$

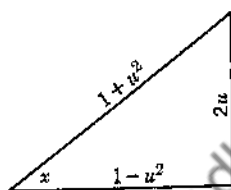


FIG. 4

This substitution thus has the desired effect. Applying it to the given integral,

$$\begin{aligned} \int \frac{1}{2 \cos x + \sin x + 3} dx &= \int \frac{1}{2 \frac{1-u^2}{1+u^2} + \frac{2u}{1+u^2} + 3} \cdot \frac{2du}{1+u^2} \\ &= \int \frac{2du}{u^2 + 2u + 5} \\ &= 2 \int \frac{du}{(u+1)^2 + 4} \\ &= \tan^{-1} \frac{u+1}{2} + C. \end{aligned}$$

In order to replace u in this answer by its value in terms of x , we must find an expression for u that satisfies both of Eqs. 10.1. Clearing fractions and transposing,

$$\begin{aligned} u^2 \sin x - 2u + \sin x &= 0, \\ u^2(1 + \cos x) - 1 + \cos x &= 0. \end{aligned}$$

Eliminating u^2 from these equations, we have

$$2u(1 + \cos x) = 2 \sin x$$

$$u = \frac{\sin x}{1 + \cos x} = \tan \frac{x}{2}.$$

Hence, the result of the above integration is

$$\tan^{-1}[(\tan x/2 + 1)/2] + C.$$

From the discussion of this example, it is clear that any integrand that is a rational function of $\sin x$ and $\cos x$ can be transformed into one that is rational in u by this transformation.

EXERCISES

Evaluate the following integrals.

$$1. \int \frac{dx}{(x+1)\sqrt{x}} = 2 \tan^{-1} \sqrt{x} + C.$$

$$2. \int \frac{dx}{\sqrt{x} + \sqrt[3]{x}}.$$

$$3. \int \frac{x dx}{(1+x)^{3/2}} = 2(\sqrt{1+x} + 1/\sqrt{1+x}) + C.$$

$$4. \int \frac{\sqrt{x} dx}{1 + \sqrt[3]{x}}.$$

$$5. \int \frac{\sqrt{4-x^2}}{x^4} dx = -(4-x^2)^{3/2}/12x^3 + C.$$

$$6. \int \frac{dx}{x(1 - \sqrt[3]{x})}.$$

$$7. \int \frac{dx}{x^3 \sqrt{1+2x+3x^2}} = \log \left| \frac{1+x+\sqrt{1+2x+3x^2}}{x} \right| - \frac{1}{x} \sqrt{1+2x+3x^2} + C.$$

$$8. \int \frac{dx}{2x\sqrt{3x^2-2x-1}}.$$

$$9. \int \frac{x^5 dx}{x^{12} + 1} = \frac{1}{6} \tan^{-1} x^6 + C.$$

$$10. \int \frac{x^7 dx}{x^{12} - 1}.$$

$$11. \int \frac{dx}{\sin x + \cos x + 1} = \log |\tan(x/2) + 1| + C.$$

$$12. \int \frac{dx}{e^x - e^{-x}}. \quad [\text{Let } x = \log u.]$$

$$13. \int \frac{dx}{4 + 5 \sin x} = \frac{1}{3} \log \left| \frac{2 \tan(x/2) + 1}{\tan(x/2) + 2} \right| + C.$$

$$14. \int \frac{\sin x \, dx}{4 \cos x + 3 \sin x}.$$

$$15. \int \frac{dx}{5 + 4 \sin 2x} = \frac{1}{3} \tan^{-1} \left(\frac{5 \tan x + 4}{3} \right) + C.$$

$$16. \int_1^5 x \sqrt{x-1} \, dx.$$

$$17. \int_0^3 \frac{\sqrt{x} \, dx}{x+1} = 2(\sqrt{3} - \pi/3).$$

$$18. \int_0^{\log \sqrt{3}} \frac{dx}{e^x + e^{-x}}.$$

$$19. \int_0^{\pi/2} \frac{d\theta}{3 + 5 \sin \theta} = \frac{1}{4} \log 3.$$

$$20. \int_0^{\pi/2} \frac{\sin x \, dx}{\sin x + \cos x + 1}.$$

11. Integral Tables. Reduction Formulas. Although it is desirable, as a matter of convenience, to be able to carry out all integrations that can be reduced to the basic formulas by the processes outlined in this chapter, some dependence upon a table of integrals is no disgrace. There are many expressions that can be reduced by one device or another to standard forms. The discovery of the appropriate artifice for any given integral, however, may be very time-consuming and of no value beyond the immediate end in view. Many results obtained by a great variety of methods, some rather obscure and difficult to find, have been collected in tables of integrals. The more extensive of these collections are published in book form,* and short collections appear in most calculus text books. A brief list of formulas will be found on pages 616-618, together with some hints for dealing with various classes of integrands.

* For example, *Tables of Integrals and Other Mathematical Data*, by H. B. Dwight, The Macmillan Co., New York. Also, integral tables in *Mathematical Tables from Handbook of Chemistry and Physics*, Chemical Rubber Publishing Co., Cleveland.

Several of the formulas given in tables of integrals are what are known as reduction formulas. Many reduction formulas are found by integration by parts. A simple example is

$$\int x^n e^x dx = x^n e^x - n \int x^{n-1} e^x dx.$$

By this formula the problem of integrating $\int x^n e^x dx$ is reduced to that of integrating a similar product with the exponent of the first factor reduced by one unit. A second application of this formula, to $\int x^{n-1} e^x dx$, gives us $\int x^n e^x dx = x^n e^x - n x^{n-1} e^x + n(n-1) \int x^{n-2} e^x dx$. By continued use of the formula one can integrate any product of the form $x^n e^x$, where n is a positive integer.

We take up one further reduction formula because of its connection with the use of partial fractions, discussed earlier in this chapter. As a preliminary step, let us integrate

$$\int \frac{1}{(x^2 + a^2)^2} dx.$$

By inspection, we write

$$\frac{1}{(x^2 + a^2)^2} = \frac{1}{a^2} \left[\frac{1}{x^2 + a^2} - \frac{x^2}{(x^2 + a^2)^2} \right].$$

Hence,

$$\begin{aligned} \int \frac{1}{(x^2 + a^2)^2} dx &= \frac{1}{a^2} \left[\int \frac{dx}{x^2 + a^2} - \int \frac{x^2 dx}{(x^2 + a^2)^2} \right] \\ &= \frac{1}{a^3} \tan^{-1} \frac{x}{a} - \frac{1}{a^2} \int \frac{x^2 dx}{(x^2 + a^2)^2}. \end{aligned}$$

The integral still to be found can be integrated by parts.

$$\begin{aligned} \int \frac{x^2 dx}{(x^2 + a^2)^2} &= \int x \cdot \frac{x dx}{(x^2 + a^2)^2} = -\frac{1}{2} \cdot \frac{x}{x^2 + a^2} + \int \frac{1}{2} \cdot \frac{dx}{x^2 + a^2} \\ &= -\frac{x}{2(x^2 + a^2)} + \frac{1}{2a} \tan^{-1} \frac{x}{a}. \end{aligned}$$

Hence,

$$\int \frac{1}{(x^2 + a^2)^2} dx = \frac{1}{2a^3} \tan^{-1} \frac{x}{a} + \frac{x}{2a^2(x^2 + a^2)} + C.$$

The same method can be applied to $\int \frac{1}{(x^2 + a^2)^n} dx$ to obtain a reduction formula. First, we note by inspection,

$$\frac{1}{(x^2 + a^2)^n} = \frac{1}{a^2} \left[\frac{1}{(x^2 + a^2)^{n-1}} - \frac{x^2}{(x^2 + a^2)^n} \right]$$

so that

$$\int \frac{1}{(x^2 + a^2)^n} dx = \frac{1}{a^2} \left[\int \frac{dx}{(x^2 + a^2)^{n-1}} - \int \frac{x^2 dx}{(x^2 + a^2)^n} \right]. \quad (11.1)$$

Integrating the second term by parts,

$$\begin{aligned} \int \frac{x^2 dx}{(x^2 + a^2)^n} &= \int x \cdot \frac{x dx}{(x^2 + a^2)^n} \\ &= -\frac{x}{2(n-1)(x^2 + a^2)^{n-1}} + \int \frac{dx}{2(n-1)(x^2 + a^2)^{n-1}}. \end{aligned}$$

Inserting this result in (11.1),

$$\begin{aligned} \int \frac{1}{(x^2 + a^2)^n} dx &= \frac{1}{a^2} \left[\int \frac{dx}{(x^2 + a^2)^{n-1}} - \frac{1}{2(n-1)} \int \frac{dx}{(x^2 + a^2)^{n-1}} \right. \\ &\quad \left. + \frac{x}{2(n-1)(x^2 + a^2)^{n-1}} \right] \\ &= \frac{1}{a^2} \left[\frac{x}{2(n-1)(x^2 + a^2)^{n-1}} + \frac{2n-3}{2n-2} \int \frac{dx}{(x^2 + a^2)^{n-1}} \right]. \end{aligned}$$

With this reduction formula, we can extend the integration of rational fractions to include those with repeated quadratic factors in the denominator.

We shall give some examples of the use of integral tables and include an example employing the above reduction formula in the method of partial fractions. Formula numbers refer to the Table of Integrals in the back of the book.

EXAMPLE 1. $\int \frac{\sqrt{7x-4}}{x} dx.$

SOLUTION. By formula 7,

$$\int \frac{\sqrt{7x-4}}{x} dx = 2\sqrt{7x-4} - 4 \tan^{-1} \frac{\sqrt{7x-4}}{2} + C.$$

EXAMPLE 2. $\int \frac{dx}{x^2\sqrt{5+2x}}.$

SOLUTION. By formula 13,

$$\int \frac{dx}{x^2\sqrt{5+2x}} = -\frac{\sqrt{5+2x}}{5x} - \frac{2}{2 \cdot 5} \int \frac{dx}{x\sqrt{5+2x}}.$$

Now, applying formula 11,

$$\int \frac{dx}{x\sqrt{5+2x}} = \frac{1}{\sqrt{5}} \log \left| \frac{\sqrt{5+2x} - \sqrt{5}}{\sqrt{5+2x} + \sqrt{5}} \right| = \frac{1}{\sqrt{5}} \log \left| \frac{5+x - \sqrt{25+10x}}{x} \right|.$$

Combining these results,

$$\int \frac{dx}{x^2\sqrt{5+2x}} = -\frac{\sqrt{5+2x}}{5x} - \frac{1}{5\sqrt{5}} \log \left| \frac{5+x - \sqrt{25+10x}}{x} \right| + C.$$

EXAMPLE 3. $\int \sqrt{\frac{x+1}{x-1}} dx.$

SOLUTION. By formula 3,

$$\begin{aligned} \int \sqrt{\frac{x+1}{x-1}} dx &= \int \frac{x+1}{\sqrt{x^2-1}} dx = \int \frac{x dx}{\sqrt{x^2-1}} + \int \frac{dx}{\sqrt{x^2-1}} \\ &= \sqrt{x^2-1} + \log |x + \sqrt{x^2-1}| + C. \end{aligned}$$

EXAMPLE 4. $\int \frac{2x^4 + 7x^3 + 15x^2 + 11x + 8}{x(x^2 + 2x + 2)^2} dx.$

SOLUTION. We write

$$\frac{2x^4 + 7x^3 + 15x^2 + 11x + 8}{x(x^2 + 2x + 2)^2} = \frac{A}{x} + \frac{Bx + C}{x^2 + 2x + 2} + \frac{Dx + E}{(x^2 + 2x + 2)^2}.$$

Evaluating the unknown coefficients in the numerators, as in Example 2, Art. 5, we obtain

$$A = 2, \quad B = 0, \quad C = -1, \quad D = 1, \quad E = -3.$$

Hence,

$$\begin{aligned} \int \frac{2x^4 + 7x^3 + 15x^2 + 11x + 8}{x(x^2 + 2x + 2)^2} dx \\ = \int \frac{2}{x} dx - \int \frac{dx}{x^2 + 2x + 2} + \int \frac{(x-3) dx}{(x^2 + 2x + 2)^2}. \end{aligned}$$

Integrating these terms separately,

$$\int 2/x dx = 2 \log |x|, \quad (11.3)$$

$$\int \frac{dx}{x^2 + 2x + 2} = \int \frac{dx}{(x+1)^2 + 1} = \tan^{-1}(x+1), \quad (11.4)$$

$$\begin{aligned}
 \int \frac{(x-3) dx}{(x^2 + 2x + 2)^2} &= \int \frac{x+1-4}{(x^2 + 2x + 2)^2} dx \\
 &= \frac{1}{2} \int \frac{(2x+2) dx}{(x^2 + 2x + 2)^2} - 4 \int \frac{dx}{(x^2 + 2x + 2)^2} \\
 &= -\frac{1}{2} \frac{1}{x^2 + 2x + 2} - 4 \int \frac{du}{(u^2 + a^2)^2}, \\
 &\quad \text{where } u = x + 1, a = 1, \\
 &= -\frac{1}{2(x^2 + 2x + 2)} - 4 \left[\frac{1}{2} \tan^{-1} u + \frac{u}{2(u^2 + 1)} \right].
 \end{aligned}$$

The last integral was found by means of the reduction formula derived in this article. Putting $u = x + 1$ in the last result, and simplifying, we have

$$\int \frac{(x-3) dx}{(x^2 + 2x + 2)^2} = -2 \tan^{-1}(x+1) - \frac{4x+5}{2(x^2 + 2x + 2)} \quad (11.5)$$

Combining the results in (11.3), (11.4), and (11.5), we have for the final answer,

$$2 \log |x| - 3 \tan^{-1}(x+1) - \frac{4x+5}{2(x^2 + 2x + 2)}.$$

EXERCISES

Integrate by means of tables or otherwise.

1. $\int \frac{x dx}{(3 + 2x)^2} = \frac{1}{4} \log |2x + 3| + \frac{3}{4(3 + 2x)} + C.$
2. $\int \sqrt{\frac{1+x}{1-x}} dx.$
3. $\int x^6 \cos x dx = x^6 \sin x + 6x^5 \cos x - 30x^4 \sin x - 120x^3 \cos x + 360x^2 \sin x + 720x \cos x - 720 \sin x + C.$
4. $\int \frac{\sin^{-1} x}{x^2} dx.$
5. $\int x^5 (\log |x|)^2 dx = \frac{1}{6} x^6 (\log |x|)^2 - \frac{1}{18} x^6 \log |x| + \frac{1}{108} x^6 + C.$
6. $\int \cos(\log |x|) dx.$
7. $\int \frac{dx}{3 + 4e^{2x}} = \frac{1}{3} x - \frac{1}{6} \log(3 + 4e^{2x}) + C.$
8. $\int \frac{ae^x + b}{ce^x + d} dx.$

$$9. \int \frac{2 dx}{(x^2 + 1)^3} = \frac{x}{2(x^2 + 1)^2} + \frac{3x}{4(x^2 + 1)} + \frac{3}{4} \tan^{-1} x + C.$$

$$10. \int \csc^8 x dx.$$

$$11. \int \sqrt{10x^2 + 7} dx = \frac{x}{2} \sqrt{10x^2 + 7} + \frac{7}{2\sqrt{10}} \log |\sqrt{10x^2 + 7} + \sqrt{10}x| + C.$$

$$12. \int \frac{dx}{(3 + 2x)\sqrt{5 + 3x}}.$$

$$13. \int \frac{dx}{\sin x \cos^5 x} = \frac{1}{5} \sec^5 x + \frac{1}{3} \sec^3 x + \sec x + \log |\csc x - \cot x| + C.$$

$$14. \int x^5 \sqrt{4 - x^4} dx.$$

$$15. \int \frac{x^4 + 8x^3 + 38x^2 + 103x + 172}{(x + 1)(x^2 + 4x + 13)^2} dx \\ = \log |x + 1| + \frac{11x + 58}{18(x^2 + 4x + 13)} + \frac{11}{54} \tan^{-1} \frac{x + 2}{3} + C.$$

12. Nonintegrable Forms. Approximations. As pointed out in the introduction to this chapter, it is easy to write a function whose integration is extremely difficult or impossible. Some examples were given there. It is, of course, useless to try to list such functions, since any list would represent a mere sampling. If one suspects that a certain integral cannot be reduced to any combination of standard formulas, he should look carefully through a fairly extensive integral table before giving up.

The term "impossible" really should be reserved for those integrals that do not exist. By the existence theorem stated in Chapter 7, the integral $\int_a^x f(t) dt$ always exists if $f(t)$ is continuous. Consequently, by the fundamental theorem, there is a function whose derivative is $f(x)$, namely, $\phi(x) = \int_a^x f(t) dt$. Hence, the integral $\int \sqrt{\sin x} dx$ exists if $0 < x < \pi$, that is, there is a function $\phi(x)$ for which $\phi'(x) = \sqrt{\sin x}$ in this range. For us, at the present time, an integration is impossible if there is no function, among all the functions that we have studied, whose derivative is the integrand. If new functions were constructed, such an integration might no longer be impossible. In fact, new functions are frequently introduced in mathematics to answer the questions posed by problems in integration. That is, the integration process is the means of constructing new functions. An important example of this genesis of a new function was studied in Chapter 11, for the function $L(x) = \int_1^x \frac{1}{t} dt$, the inverse

derivative of $1/x$. One of the most important uses of integration is that of generating new functions to meet the needs of various situations in science.

If some given integral is indeed impossible of inverse differentiation by elementary formulas, there are methods by which it may be possible to approximate it. One of these methods will be discussed in Chapter 20. Definite integrals, which we usually find by means of inverse differentiation, can be approximated by several methods, some of which we shall study in Chapter 19.

MISCELLANEOUS EXERCISES IN INTEGRATION

Integrate by any method.

$$1. \int \frac{\sin \sqrt{x}}{\sqrt{x}} dx.$$

$$3. \int \frac{e^{2x} - 1}{e^{2x} + 1} dx.$$

$$5. \int \frac{dx}{x^2 + 6x + 5}.$$

$$7. \int \tan^{-1} \sqrt{x} dx.$$

$$9. \int (\sin^2 \theta + 1)^2 d\theta.$$

$$11. \int \frac{t dt}{\sqrt{9 - t^2}}.$$

$$13. \int \sin 4x \sin 2x dx.$$

$$15. \int x^2 \sin^{-1} x dx.$$

$$17. \int \frac{x^3 + 1}{x(x-1)^2} dx.$$

$$19. \int \frac{dx}{x(1 + \sqrt{x})}.$$

$$21. \int \frac{e^{2t} dt}{e^t + e^{-t}}.$$

$$23. \int x^3 \sin^2 x dx.$$

$$25. \int e^t \sin^2 t dt.$$

$$2. \int \frac{\log \sqrt{x}}{x} dx.$$

$$4. \int \frac{dx}{1 - \cos x}.$$

$$6. \int \frac{\sqrt{25 - x^2}}{x^3} dx.$$

$$8. \int \frac{xe^x}{(1+x)^2} dx.$$

$$10. \int \frac{4 + 3x}{x^2 + x + 1} dx.$$

$$12. \int \sqrt{4 - 2u - u^2} du.$$

$$14. \int \sin^4 \theta \cos^4 \theta d\theta.$$

$$16. \int e^{2t} \sin \pi t dt.$$

$$18. \int \frac{x^5 dx}{(x^2 + 1)^2}.$$

$$20. \int \frac{dx}{x\sqrt{x^2 - x + 1}}.$$

$$22. \int \frac{dt}{2 - \cos 2t} \quad [\text{Put } 2t = s.]$$

$$24. \int \frac{\cos t dt}{3 \sin t + 4 \sin^2 t}.$$

$$26. \int \frac{dx}{x^4 \sqrt{4 - x^2}}.$$

$$27. \int \frac{dx}{\sqrt{(1+x)(5-x)}}$$

$$29. \int \frac{(9x^2 + 4)^{\frac{3}{2}} dx}{x}$$

$$31. \int_0^a (\sqrt{x} + \sqrt{a})^2 dx.$$

$$33. \int_0^2 \frac{3x dx}{x^4 + 4}$$

$$35. \int_0^{\frac{1}{2}} \sqrt{1-4x^2} dx.$$

$$37. \int_2^5 \frac{3v^2 + 5v}{(v-1)(v+1)^2} dv.$$

$$39. \int_{64}^{729} \frac{dt}{\sqrt{t} - \sqrt[3]{t}}$$

$$28. \int t^2 \sqrt{4-9t^2} dt. \quad \left[\text{Put } u = \sin^{-1} \frac{3t}{2} \right].$$

$$30. \int \sqrt{\frac{4+3t}{4-3t}} dt.$$

$$32. \int_0^{\log 4} \sqrt{e^t} dt.$$

$$34. \int_8^{12} \frac{du}{u^2 - 4u}$$

$$36. \int_3^4 \frac{dt}{t\sqrt{25-t^2}}$$

$$38. \int_1^{16} \frac{x^{\frac{1}{2}} dx}{1+x^{\frac{1}{2}}}$$

$$40. \int_0^{\pi/2} \frac{d\theta}{2 + \sin \theta}$$

Geometrical Applications Rectangular Coordinates

1. The Tangent Problem and Curve Tracing. It has been apparent from the first that the solution of the tangent problem could be of great assistance in drawing graphs. Some use of the derivative in curve-tracing was made in earlier chapters. Here we shall try to exploit this use to a greater degree.

By means of the derivative we can find the direction of a curve at doubtful points. Consider, for example, the graph $y = e^{-x^2}$. We find that the curve passes through $(0, 1)$, which is its highest point, since y

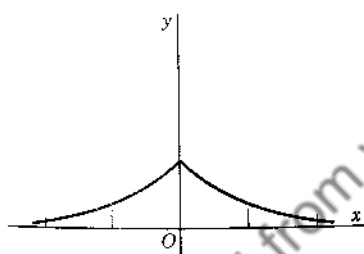


FIG. 1

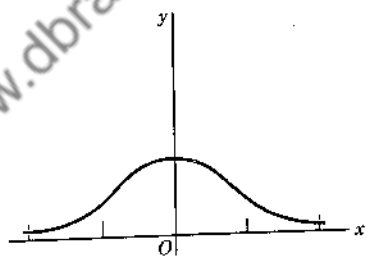


FIG. 2

must decrease as x becomes numerically greater. The x -axis is an asymptote in both directions since y cannot become 0 but must approach 0 as x increases indefinitely. Hence, we know that the curve looks like Fig. 1 or Fig. 2, but we cannot be sure which is correct without finding the slope at $x = 0$. We could draw the same conclusions about the graph of $y = e^{-|x|}$. Do these equations have graphs like those shown in Figs. 1 and 2, and, if so, which graph is associated with one equation and which with the other?

Differentiating

$$y = e^{-x^2},$$

we have

$$y' = -2xe^{-x^2}.$$

Hence, at $x = 0$,

$$y' = 0,$$

and the graph has the appearance of Fig. 2.

The function $y = e^{-|x|}$ is identical with

$$y = e^{-x} \text{ for } x \geq 0, \quad y = e^x \text{ for } x \leq 0.$$

Differentiating,

$$y' = -e^{-x} \text{ for } x \geq 0, \quad y' = e^x \text{ for } x \leq 0.$$

At $x = 0$, the function $e^{-|x|}$ cannot be said to have a derivative. As $x \rightarrow 0$, however, the derivative approaches -1 if the approach is from the positive side of 0 and $+1$, if from the negative side. Hence, the curve $y = e^{-|x|}$ has a peak at $x = 0$, as shown in Fig. 1.

Other uses of the derivative in curve-tracing include the location of maximum and minimum points and points of inflection. These applications will be discussed in Arts. 3 and 4.

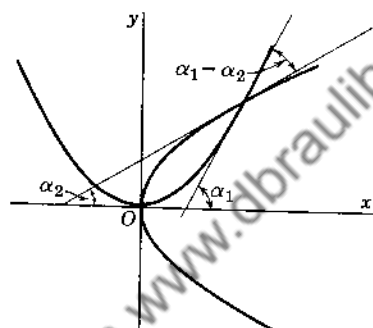


FIG. 3

2. Intersections of Two Curves. The angles of intersection of two curves are often a matter of interest. For example, the parabolas $y = x^2$ and $y = \sqrt{x}$ intersect in points where

$$x^2 = \sqrt{x},$$

namely, $x = 0$ and $x = 1$. The points of intersection are thus $(0, 0)$ and $(1, 1)$ (Fig. 3). Differentiating the equations of these parabolas, we obtain

$$y' = 2x \quad \text{and} \quad y' = \frac{1}{2\sqrt{x}}, \quad \text{respectively.}$$

At the point $(0, 0)$, the slope of the parabola $y = x^2$ is 0, whereas that of the parabola $y = \sqrt{x}$ is infinite. That is, the first curve has a horizontal tangent at this point and the second a vertical tangent. Hence, at this point the curves intersect at right angles.

At the point $(1, 1)$, the slopes are, respectively,

$$m_1 = \tan \alpha_1 = 2 \quad \text{and} \quad m_2 = \tan \alpha_2 = \frac{1}{2},$$

where α_1 and α_2 are the angles of inclination of the tangents. Since the angle between the tangents is $\alpha_1 - \alpha_2$, we have

$$\tan(\alpha_1 - \alpha_2) = \frac{\tan \alpha_1 - \tan \alpha_2}{1 + \tan \alpha_1 \tan \alpha_2} = \frac{m_1 - m_2}{1 + m_1 m_2} = \frac{2 - 1/2}{1 + 2 \cdot 1/2} = \frac{3}{4}$$

Hence, $\alpha_1 - \alpha_2 = \tan^{-1} \frac{3}{4} = 36^\circ 52'$, approximately.

In general, if the curves $y = f(x)$ and $y = g(x)$ intersect at (x_1, y_1) ,

$$\tan(\alpha_1 - \alpha_2) = \frac{f'(x_1) - g'(x_1)}{1 + f'(x_1)g'(x_1)} \quad (2.1)$$

If the curves intersect at right angles at this point,

$$1 + f'(x_1)g'(x_1) = 0, \quad \text{or} \quad f'(x_1) = -\frac{1}{g'(x_1)} \quad (2.2)$$

EXERCISES

Show that each of the curves in Exercises 1-6 passes through the point $(0, 1)$ and find its slope there.

1. $y = 1/(1 + x^2)$.
2. $y = (1 + x)/(1 + x^2)$.
3. $y = (1 - x^2)/(x - 1)$.
4. $y = (1 + x)e^{-x}$.
5. $y = 1 + \tan^{-1} x$.
6. $y = \tan(x + \pi/4)$.

Find the angles of intersection of the pairs of curves in Exercises 7-12.

7. $xy = 1$, $y = \sqrt{x}$.
8. $y = x^3$, $y = \sqrt[3]{x}$.
9. $y = 1/x$, $y = 2/(1 + x^2)$.
10. $y = \sin x$, $y = \cos x$.
11. $y = e^{-x}$, $y = e^{-x^2}$.
12. $y = \sin^{-1} x$, $y = \frac{\pi}{2} x$.

13. Show that the ellipse $\frac{x^2}{625} + \frac{y^2}{400} = 1$ intersects the hyperbola $\frac{x^2}{144} - \frac{y^2}{81} = 1$ at right angles.

14. Find the angle of intersection of the curves $x^2 + y^2 - 8x = 0$ and $y^2 = x^3/(2 - x)$ at a point where $x = 8/5$.

15. Find the angle at which the parabola $y = 11 - x^2$ intersects the hyperbola $y^2 - x^2 = 1$ in the first quadrant.

3. Maximum, Minimum, and Inflection Points. Tests. We saw in Chapter 5 how maximum and minimum points can be located, by setting the derived function equal to 0. In complicated cases it is not as easy as we found it there to determine the nature of the critical points resulting from this equation. We now consider this problem again with more difficult examples, and shall devise tests to decide the type of critical point found.

EXAMPLE 1. Locate maxima, minima, and inflection points and draw the graph of

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

SOLUTION. Differentiating,

$$\begin{aligned} y' &= 2x^4 - 7x^3 + 3x^2 + 7x - 5 \\ &= (x+1)(x-1)^2(2x-5). \end{aligned}$$

Hence, $y' = 0$ at $x = -1$, $x = 1$, and $x = 5/2$. By substitution in the expression for y , we find the critical points to be

$$(-1, 87/20), (1, -57/20), \text{ and } (5/2, -339/64).$$

We shall consider two ways by which the type of each of these critical points can be determined. Each of these ways has its advantages and both have general applicability. The first of these consists of a comparison of the slope on the two sides of the critical point. This is easy in the present example, since the derived function is expressed in factored form. For this test it is best to examine the critical points in order of their occurrence from left to right on the graph, and arrange the factors of the derived function in the same order as the corresponding critical points. Thus we write, as above,

$$y' = (x+1)(x-1)^2(2x-5),$$

the critical points occurring for $x = -1$, $x = 1$, and $x = 5/2$, in order from left to right.

Now, to the left of $x = -1$ all factors of y' are negative, and since there are four of them the product is positive. Hence, to the left of -1 the slope is positive. Between -1 and $+1$, that is, $-1 < x < +1$, the first factor is positive whereas the other three are negative. Hence, the product is negative, and the slope is negative in this interval. Thus, the directions of the tangents near this point are arranged as shown in the diagram, , making it a maximum point. All this may be shown concisely, not only for this point but for the other critical points, as follows.

Point	$y' = (x+1)(x-1)^2(2x-5)$	Slope	Type
$(-1, \frac{87}{20});$	$\begin{cases} x < -1 \\ -1 < x < 1 \end{cases}$	$\begin{matrix} - & - & - \\ + & - & - \end{matrix}$	max
$(1, -\frac{57}{20});$	$\begin{cases} -1 < x < 1 \\ 1 < x < \frac{5}{2} \end{cases}$	$\begin{matrix} + & - & - \\ + & + & - \end{matrix}$	level inflection
$(\frac{5}{2}, -\frac{339}{64});$	$\begin{cases} 1 < x < \frac{5}{2} \\ \frac{5}{2} < x \end{cases}$	$\begin{matrix} + & + & - \\ + & + & + \end{matrix}$	min

The graph is shown in Fig. 4.

The second way of classifying the critical points makes use of the second derivative. From Chapter 4 we recall that the curve is concave upward at a point where $y'' > 0$, concave downward at a point where $y'' < 0$, whereas at an inflection point $y'' = 0$. Note that we do not say that an inflection point occurs wherever $y'' = 0$, and, in fact, examples were given in Chapter 4 in which a minimum occurred even though $y'' = 0$. Therefore, the second derivative test is not always conclusive.

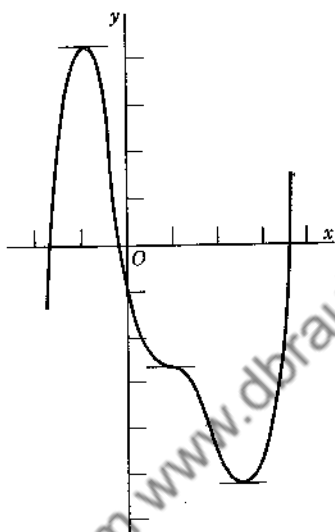


FIG. 4

It is nevertheless useful, and often easier to apply than the test described above. Let us see how it works in the present example. Differentiating again,

$$y'' = 8x^3 - 21x^2 + 6x + 7.$$

We must now find the sign of y'' at each critical point by substituting the value of x corresponding to the critical point in this expression.

At $x = -1$, $y'' = -8 - 21 - 6 + 7 < 0$, and the point is a maximum.

At $x = 1$, $y'' = 8 - 21 + 6 + 7 = 0$, and the point *may be* an inflection point.

At $x = 5/2$, $y'' = 125 - 525/4 + 15 + 7 > 0$, and the point is a minimum.

We are left in doubt about $x = 1$, and must decide whether or not it is really an inflection point by other means. We can test this point by a

method similar to the first derivative test as follows: for x a little less than 1 we find $y'' > 0$, and, hence, the direction of concavity is upward; for x a little greater than 1 we find $y'' < 0$, and, hence, the direction of concavity is downward; since the point $x = 1$ is a point where the direction of concavity changes, it is an inflection point.

These two tests can be summarized as follows.

TEST BY THE FIRST DERIVATIVE

1. Consider critical points in order from left to right on the graph. If the first derivative is factored, arrange the factors in the same order.
2. Determine the sign of the derivative in each interval between critical points.
3. It is now possible to determine the nature of each critical point by inspection of these signs. If, as the curve proceeds through a critical point, the values of the derivative are $- 0 +$ in order, the point is a minimum; if $+ 0 -$, a maximum; if $+ 0 +$ or $- 0 -$, a level inflection point.

TEST BY THE SECOND DERIVATIVE

1. Substitute the coordinates (often only x) of the critical point in the second derivative.
2. If the second derivative is negative at a critical point, this point is a maximum; if it is positive, the point is a minimum; if it is 0, this test is inconclusive and the point should be tested by means of the first derivative.

EXAMPLE 2. Examine for critical points and draw the graph:

$$y = x e^{-x}.$$

SOLUTION. Differentiating,

$$y' = (1 - x)e^{-x}.$$

$$y'' = (x - 2)e^{-x}.$$

The only point of 0 slope is $x = 1$, $y = e^{-1} = 1/e = .37$, approximately. At this point, $y'' = (1 - 2)e^{-1} < 0$, and hence the point is a maximum point.

The second derivative is equal to 0 at $x = 2$, $y = 2e^{-2} = .27$, approximately, and hence this point may be an inflection point. Since the curve is concave downward for values of x that are less than 2 and concave upward for values that are greater than 2, it is, in fact, an inflection point.

In Chapter 11, Art. 14, we found that $xe^{-x} \rightarrow 0$ as $x \rightarrow \infty$. Therefore, the curve approaches the positive x -axis asymptotically. As $x \rightarrow -\infty$,

$e^{-x} \rightarrow \infty$, and hence $xe^{-x} \rightarrow -\infty$. The curve passes through the origin with its slope there equal to 1. The graph is shown in Fig. 5.

EXERCISES

Examine for maximum and minimum points and points of inflection, and draw the graph of each of the following equations.

1. $y = x^4 + 4x^3 - 16x - 8.$

2. $y = 2x^3 - 3x^2 - 12x - 6.$

3. $y = (x^2 + x + 4)/(x + 1).$

4. $y = (1 + x)^2(2 - x)^2.$

5. $y = x^2e^{-x}.$

6. $y = xe^{-x^2}.$

7. $y = \log \tan x, \quad 0 < x < \pi/2.$

8. $y = x \log x \quad \text{for } x > 0,$
 $= 0 \quad \text{for } x = 0.$

9. $y = x/2 - \sin x, \quad 0 < x < 2\pi.$

10. $y = e^{-x} \sin x.$

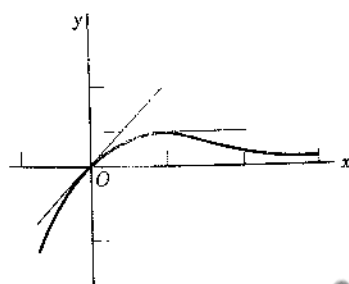


FIG. 5

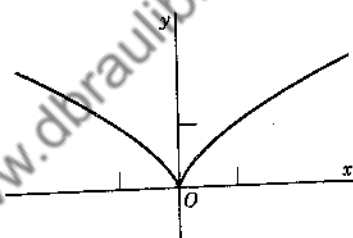


FIG. 6

4. Irrational Algebraic Functions. For the most part, we have thus far dealt with functions for which maximum and minimum points are associated with horizontal tangents. Graphs of irrational functions often have maximum and minimum points associated with vertical tangents. At such points, the derivative does not exist and the derived function is discontinuous. We shall consider some examples.

EXAMPLE 1. $y = x^{\frac{2}{3}}.$

SOLUTION. Differentiating,

$$y' = \frac{2}{3} x^{-\frac{1}{3}} = \frac{2}{3\sqrt[3]{x}}.$$

Clearly, there is no point for which $y' = 0$. As x approaches 0, however, the numerical value of the slope approaches ∞ , and the tangent tends to a vertical position. From the equation, $y = x^{\frac{2}{3}}$, it is clear that y is always positive, except for $x = 0$, where it has its minimum value. The graph is shown in Fig. 6.

EXAMPLE 2. Examine for maximum and minimum points and points of inflection and draw the graph of $y = 12x^{\frac{2}{3}} - x$.

SOLUTION. Differentiating,

$$y' = 4x^{-\frac{1}{3}} - 1 = \frac{4}{x^{\frac{1}{3}}} - 1.$$

$$y'' = -\frac{8}{3x^{\frac{4}{3}}}.$$

Setting $y' = 0$, we have $x^{\frac{1}{3}} = 4$, and $x = \pm 8$. Substituting these values in the given equation for y , we find that the critical points are $x = 8$, $y = 16$, and $x = -8$, $y = -16$. Now, at $x = 8$, y'' is negative, and hence $(8, 16)$ is a maximum point; and at $x = -8$, y'' is positive, and hence $(-8, -16)$ is a minimum point. There is no value of x for which $y'' = 0$. It is nevertheless true that $(0, 0)$ is an inflection point, since the

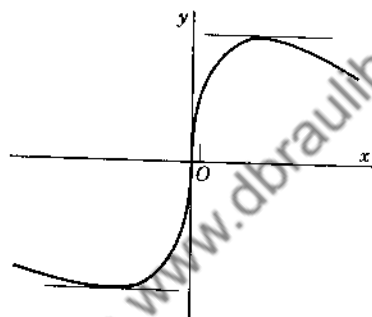


FIG. 7

curve is concave upward for $x < 0$ and concave downward for $x > 0$. This illustrates the fact that the rule that $y'' = 0$ at inflection points may fail at points of discontinuity of y' or y'' . At $(0, 0)$, both of these derived functions are discontinuous. The curve has a vertical tangent at this point. The graph, shown in Fig. 7, would meet the x -axis at $x = \pm 12\sqrt{12}$, if extended.

EXAMPLE 3. Examine for maximum and minimum points and points of inflection and draw the graph of

$$y = (2 - x)^{\frac{1}{3}}(1 - x)^{\frac{2}{3}}.$$

SOLUTION. Differentiating,

$$\begin{aligned} y' &= -\frac{1}{3}(2 - x)^{-\frac{2}{3}}(1 - x)^{\frac{2}{3}} - \frac{2}{3}(2 - x)^{\frac{1}{3}}(1 - x)^{-\frac{1}{3}} \\ &= -\frac{1}{3} \frac{1}{(2 - x)^{\frac{2}{3}}(1 - x)^{\frac{1}{3}}} [(1 - x) + 2(2 - x)] \\ &= \frac{3x - 5}{3(2 - x)^{\frac{2}{3}}(1 - x)^{\frac{1}{3}}}. \end{aligned}$$

$$\begin{aligned}
 y'' &= \frac{1}{3}[3(2-x)^{-\frac{2}{3}} \cdot (1-x)^{-\frac{1}{3}} + (3x-5) \cdot \frac{2}{3}(2-x)^{-\frac{5}{3}}(1-x)^{-\frac{1}{3}} \\
 &\quad + (3x-5)(2-x)^{-\frac{1}{3}} \cdot \frac{1}{3}(1-x)^{-\frac{4}{3}}] \\
 &= \frac{1}{9} \frac{1}{(2-x)(1-x)} [9(2-x)(1-x) + 2(3x-5)(1-x) \\
 &\quad + (3x-5)(2-x)] \\
 &= -\frac{2}{9(2-x)^{\frac{5}{3}}(1-x)^{\frac{5}{3}}}.
 \end{aligned}$$

From the expressions for y , y' , and y'' we obtain the following information.

$$y = 0 \text{ for } x = 1 \text{ and } x = 2.$$

$$y' = 0 \text{ for } x = 5/3, y = \sqrt[3]{4/3}.$$

At this point, $y'' < 0$. Hence, it is a maximum point.

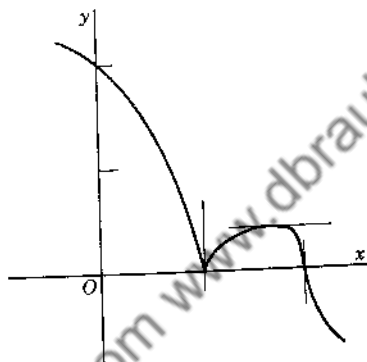


FIG. 8

y'' is never 0. But, since y'' changes sign from minus to plus at $x = 2$, this is an inflection point. The tangent is vertical at this point since y' tends to $-\infty$ as x tends to 2.

y' also tends to $\pm \infty$ as x tends to 1. The point (1, 0) is a minimum point since $y > 0$ for values of x less than 1 or lying between 1 and 2. The graph is shown in Fig. 8.

EXERCISES

Examine for maximum and minimum points and points of inflection and draw the graphs.

1. $y = x^{\frac{1}{3}}$.

2. $y = (x-1)^{\frac{2}{3}}$.

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

4. $y = 2x - (x-8)^{\frac{2}{3}}$.

5. $y = (x-1)x^{\frac{2}{3}}$.

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

5. Curves Given by Parametric Equations. For the purpose of studying curves it is often advantageous to have the coordinates x and y expressed

in terms of a third variable, or parameter. We shall consider some examples.

EXAMPLE 1. The coordinates of the points on the circle $x^2 + y^2 = 25$ can be given by the parametric equations

$$x = 5 \cos t \quad \text{and} \quad y = 5 \sin t.$$

The parameter t clearly represents the angle between the radius vector to any point (x, y) and the x -axis (Fig. 9). Let us find the slope of the tangent at any point from the parametric equations. Differentiating,

$$dx/dt = -5 \sin t, \quad dy/dt = 5 \cos t.$$

Dividing, $dy/dx = (dy/dt)/(dx/dt) = -\cot t$. This result should have been obvious from the start, since the slope of the normal at (x, y) , which is the radius at that point, is $y/x = \tan t$.

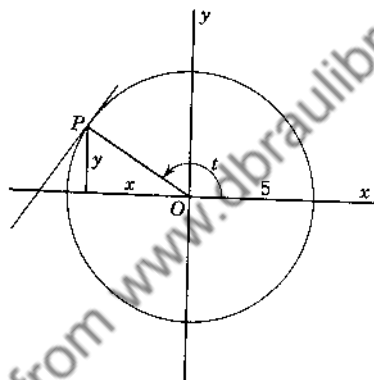


FIG. 9

EXAMPLE 2. As a more difficult example let us consider the curve whose parametric equations are

$$x = 2 \cos \theta - \cos 2\theta - 1$$

$$y = 2 \sin \theta - \sin 2\theta.$$

Differentiating,

$$dy/d\theta = 2 \cos \theta - 2 \cos 2\theta,$$

$$dx/d\theta = -2 \sin \theta + 2 \sin 2\theta,$$

whence

$$\begin{aligned} dy/dx &= \frac{\cos \theta - \cos 2\theta}{-\sin \theta + \sin 2\theta} \\ &= \frac{\cos \theta - 2 \cos^2 \theta + 1}{-\sin \theta + 2 \sin \theta \cos \theta} \\ &= \frac{(1 - \cos \theta)(2 \cos \theta + 1)}{\sin \theta(2 \cos \theta - 1)}. \end{aligned}$$

The numerator of this fraction is 0 for

$$\theta = 0, \quad (x = 0, y = 0),$$

$$\theta = 2\pi/3, \quad (x = -3/2, y = 3\sqrt{3}/2),$$

and

$$\theta = 4\pi/3, \quad (x = -3/2, y = -3\sqrt{3}/2).$$

The denominator is 0 for

$$\theta = 0, \quad (x = 0, y = 0),$$

$$\theta = \pi, \quad (x = -4, y = 0),$$

$$\theta = \pi/3, \quad (x = 1/2, y = \sqrt{3}/2),$$

$$\theta = 5\pi/3, \quad (x = 1/2, y = -\sqrt{3}/2).$$

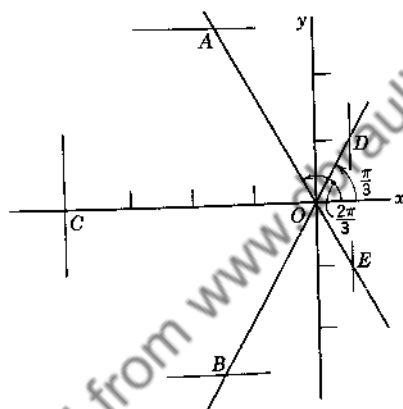


FIG. 10

Clearly, $dy/dx = 0$ for $\theta = 2\pi/3$ and $4\pi/3$, so that $A(-3/2, 3\sqrt{3}/2)$ and $B(-3/2, -3\sqrt{3}/2)$ are points of horizontal tangency. Since dy/dx is infinite at $\theta = \pi$, $\theta = \pi/3$, and $\theta = 5\pi/3$, the points $C(-4, 0)$, $D(1/2, \sqrt{3}/2)$, and $E(1/2, -\sqrt{3}/2)$ are points of vertical tangency (Fig. 10).

At $\theta = 0$, corresponding to the point $O(0, 0)$, both numerator and denominator of the expression for dy/dx vanish, and hence this function is discontinuous at 0. In this contingency, we can consider the value of the slope for a point near $(0, 0)$, and examine its behavior as the point tends to $(0, 0)$, i.e., as $\theta \rightarrow 0$. We write $\frac{dy}{dx} = \frac{1 - \cos \theta}{\sin \theta} \cdot \frac{2 \cos \theta + 1}{2 \cos \theta - 1}$,

and observe that as $\theta \rightarrow 0$ the second fraction tends to 3, whereas the first fraction tends to 0. Since the curve is continuous here, both x and y being continuous functions of θ , we conclude that at this point the tangent is horizontal.

In addition to the information obtained from the first derivative, we find the second derivative in order to determine the direction of concavity. We have

$$\begin{aligned}\frac{d^2y}{dx^2} &= \frac{d}{dx} \left(\frac{dy}{dx} \right) = \frac{\frac{d}{d\theta} \left(\frac{dy}{dx} \right)}{\frac{dx}{d\theta}} = \frac{\frac{d}{d\theta} \left(\frac{\cos \theta - \cos 2\theta}{\sin 2\theta - \sin \theta} \right)}{2 \sin 2\theta - 2 \sin \theta} \\ &= \frac{3(1 - \cos \theta)}{\sin^3 \theta (2 \cos \theta - 1)^3}.\end{aligned}$$

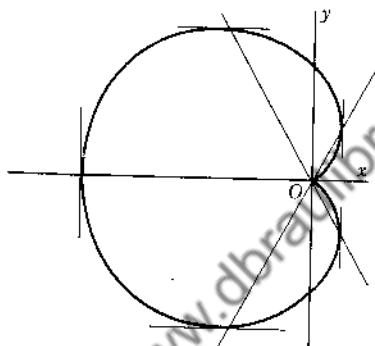


FIG. 11

By examining the factors in this fraction, we find that

for $0 < \theta < \pi/3$ the signs are $\frac{+}{++}$, hence concave up;

for $\pi/3 < \theta < \pi$ the signs are $\frac{+}{+-}$, hence concave down;

for $\pi < \theta < 5\pi/3$ the signs are $\frac{+}{--}$, hence concave up;

for $5\pi/3 < \theta < 2\pi$ the signs are $\frac{+}{-+}$, hence concave down.

It should be observed from the equations for x and y that x assumes the same value for $-\theta$ as for $+\theta$, whereas the value of y corresponding to $-\theta$ is the negative of its value for $+\theta$. This means that the part of the curve below the x -axis is symmetric to that above the x -axis. Figure 11 shows the complete curve. It is called a *cardioid*.

EXAMPLE 3. Let a curve be generated as follows. A circle of radius a is tangent to the x -axis at the origin, the center being $(0, a)$. This circle

is rolled along the positive x -axis, and as it does so the point P of the circumference, which is initially at the origin, describes a curve. Let us find the equations of this moving point in terms of the angle of rotation, θ , as a parameter. Using the equations so found, we can then find points of horizontal and vertical tangency, and the equations of the tangent and normal at any point.

From Fig. 12 we note that arc $AP = a\theta = OA$. For x , we therefore have

$$x = OA - MK = a\theta - a \sin \theta = a(\theta - \sin \theta).$$

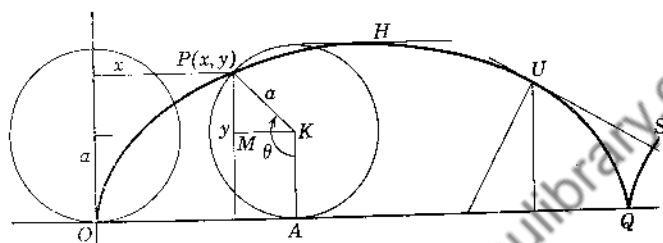


FIG. 12

Similarly, for y ,

$$y = AK + MP = a - a \cos \theta = a(1 - \cos \theta).$$

These are parametric equations of the curve $OPHQ$ S, which is called a *cycloid*.

Differentiating these equations,

$$dx/d\theta = a(1 - \cos \theta),$$

$$dy/d\theta = a \sin \theta,$$

and

$$dy/dx = \sin \theta / (1 - \cos \theta).$$

Since $\sin \theta = 0$ at $\theta = 0, \pi, 2\pi$, etc., and $1 - \cos \theta = 0$ at $\theta = 0, 2\pi, 4\pi$, etc., we see that at $\theta = \pi, 3\pi$, etc., there are points of horizontal tangency. At $\theta = 0, 2\pi$, etc., the tangency is vertical, since as θ tends to $0, 2\pi$, or any even multiple of π , the fraction $\sin \theta / (1 - \cos \theta)$ becomes infinite.

Using the first derivative test, we find that the points of horizontal tangency are all maximum points. The denominator of $\sin \theta / (1 - \cos \theta)$ is never negative, whereas the sign of the numerator changes from positive to negative as θ passes through any odd multiple of π . The student can show that the direction of concavity is downward except at $\theta =$ an even multiple of π , where both the first and second derivatives are discontinuous and the curve has *cusps*.

From the expression for dy/dx we can write the equations of the tangent and normal at any point. Let us find these equations for the point U , where $\theta = 4\pi/3$, $x = (4\pi/3 + \sqrt{3}/2)a$, $y = (1 + \frac{1}{2})a = \frac{3}{2}a$. The slope of the tangent here is $-\sqrt{3}/3$. The equation of the tangent is thus

$$y - \frac{3}{2}a = -\frac{\sqrt{3}}{3} \left[x - \left(\frac{4\pi}{3} + \frac{\sqrt{3}}{2} \right) a \right]$$

or

$$7x + 12y = 53a, \text{ approximately.}$$

The equation of the normal is

$$y - \frac{3}{2}a = \sqrt{3} \left[x - \left(\frac{4\pi}{3} + \frac{\sqrt{3}}{2} \right) a \right]$$

or

$$26x - 15y = 109a, \text{ approximately.}$$

EXERCISES

Find the equations of the tangent and normal of each of the curves in Exercises 1-10 at the point specified.

1. The ellipse, $x = 4 \cos \phi$, $y = 3 \sin \phi$; $\phi = \pi/4$.
2. The parabola, $x = 100t$, $y = -16t^2$; $t = 2$.
3. The circle, $x - 5 = 3 \cos \theta$, $y + 2 = 3 \sin \theta$; $\theta = \pi/3$.
4. The hyperbola, $x = 3t$, $y = 4/t$; $x = 3$, $y = 4$.
5. The hyperbola, $x = 2e^t$, $y = 5e^{-t}$; $t = 0$.
6. The hyperbola, $x = (1 + t^2)/(1 - t^2)$, $y = 2t/(1 - t^2)$; $t = 2$.
7. The ellipse, $x = 3(2t)/(1 + t^2)$, $y = (1 - t^2)/(1 + t^2)$; $t = 2$.
8. The hypocycloid, $x = 4 \cos^3 \theta$, $y = 4 \sin^3 \theta$; $\theta = \pi/6$.
9. The parabolic arc, $x = a \cos^4 \theta$, $y = a \sin^4 \theta$; $\theta = \pi/4$.
10. The hyperbola, $x = 2 + 3 \tan \theta$, $y = 1 + 4 \sec \theta$; $x = \frac{11}{4}$, $y = \frac{13}{4}$.

Find the points of horizontal and vertical tangency, intervals of upward and downward concavity, and draw the curves.

11. $x = 4t^3$, $y = 2t^2$.
12. $x = t^3$, $y = 2t$.
13. $x = \log t$, $y = 1/t$; $t > 0$.
14. $x = \sin 2\theta$, $y = \cos \theta$.
15. $x = 2(\theta + \sin \theta)$, $y = 2(1 - \cos \theta)$.
16. $x = \sin^2 \theta$, $y = \sin^2 \theta \tan \theta$.
17. $x = e^{-t}$, $y = \sqrt{t}$; $t \geq 0$. [Note that $dx/dy = 0$ at a point for which $dx/dt \neq 0$.]
18. $x = 3t/(1 + t^3)$, $y = 3t^2/(1 + t^3)$.

6. Length of a Curve. In Art. 7, Chapter 6, the distance a body moves along a curve was introduced through v , its velocity along the curve, which was conceived as the derivative of the distance s , i.e., ds/dt . It was then possible to find s by inverse differentiation, although we were able to do this only in the simplest examples, owing to lack of technique of integration at that time. We now derive the formula for distance by a different approach, in the form of a definite integral.

Let us consider the curve whose equation is $4y = x^2$. We set ourselves the problem of finding the length of the arc joining $P(1, 1/4)$ and $Q(3, 9/4)$ (Fig. 13). The elementary definition of length, designed for

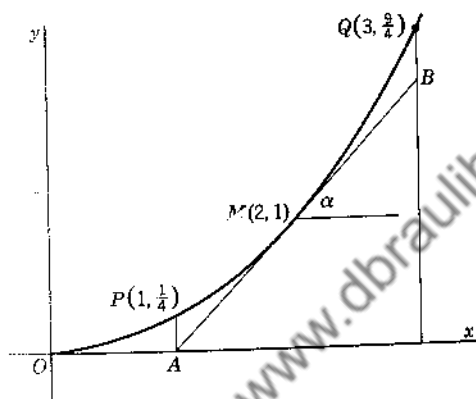


FIG. 13

application to straight-line segments, cannot be used here. But we can employ this definition to obtain approximations to the length of a curve just as we obtained approximations to the area bounded by a curve by means of the elementary definition of area. We draw a tangent to the arc at some intermediate point, say $M(2, 1)$, and take as a first approximation the length AB of this tangent intercepted by the ordinates $x = 1$ and $x = 3$ (Fig. 13). Differentiating the equation of the curve, we have $y' = x/2$, whence the slope of the tangent at $(2, 1)$ is $\tan \alpha = 2/2 = 1$. Therefore, the length of AB is

$$AB = 2 \sec \alpha = 2\sqrt{1 + \tan^2 \alpha} = 2\sqrt{2}.$$

To obtain a second approximation, we might subdivide the interval $1 \leq x \leq 3$ by the points $1, 3/2, 2, 5/2, 3$; draw tangents in each subinterval; and sum their lengths (Fig. 14). If we draw tangents at $x = 5/4, 7/4, 9/4$, and $11/4$, their slopes are, respectively,

$$\tan \alpha_1 = 5/8, \quad \tan \alpha_2 = 7/8, \quad \tan \alpha_3 = 9/8, \quad \tan \alpha_4 = 11/8.$$

The sum of the lengths of the segments of these tangents intercepted by ordinates drawn at the points of subdivision of the interval is

$$\begin{aligned} & \frac{1}{2}[\sec \alpha_1 + \sec \alpha_2 + \sec \alpha_3 + \sec \alpha_4] \\ &= \frac{1}{2}[\sqrt{1 + (5/8)^2} + \sqrt{1 + (7/8)^2} + \sqrt{1 + (9/8)^2} + \sqrt{1 + (11/8)^2}]. \end{aligned}$$

In this way, we form a sequence of approximations to the curve, each made up of the tangents at intermediate points of the subintervals. If

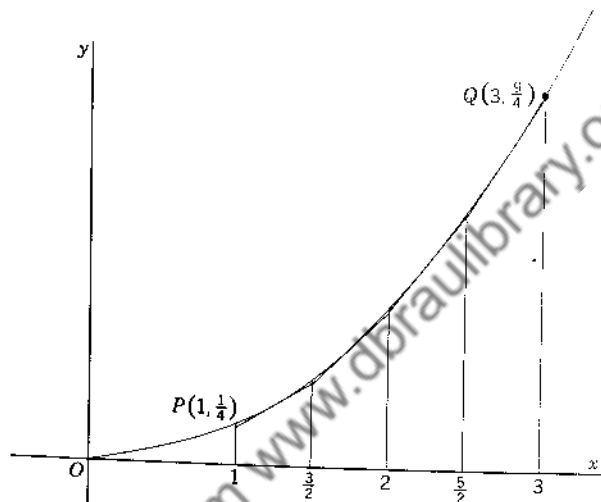


FIG. 14

the number of subintervals is n , the sum of the lengths of the tangents is

$$\begin{aligned} \sum_{i=1}^n \sec \alpha_i \cdot \Delta_i x &= \sum_{i=1}^n \sqrt{1 + \tan^2 \alpha_i} \Delta_i x \\ &= \sum_{i=1}^n \sqrt{1 + (x_i/2)^2} \Delta_i x, \end{aligned}$$

where x_i is the abscissa of the point at which the tangent is drawn in the i th subinterval and $\Delta_i x$ is the length of the subinterval. Clearly, the set of tangents approximates the curve more and more closely as n increases and the length of each subinterval decreases. We define the length of the curve to be the limit of such a sequence of approximations. That is,

$$\begin{aligned} \text{Length} = s &= \lim_{n \rightarrow \infty} \sum_{i=1}^n \sqrt{1 + (x_i/2)^2} \Delta_i x = \int_1^3 \sqrt{1 + (x/2)^2} dx \\ &= (3\sqrt{13} - \sqrt{5})/4 + \log(3 + \sqrt{13})/(1 + \sqrt{5}). \end{aligned}$$

This method can be applied to find the length of any curve $y = f(x)$, between the points (a, c) and (b, d) , if $f'(x)$ is continuous in this interval. The general formulation is exactly like that in the above example, and leads to

$$s = \lim_{n \rightarrow \infty} \sum_{i=1}^n \sqrt{1 + [f'(x_i)]^2} \Delta x = \int_a^b \sqrt{1 + [f'(x)]^2} dx,$$

or

$$s = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx. \quad (6.1)$$

If the curve is given by the equation $x = g(y)$, the length of the arc between (a, c) and (b, d) , provided that $g'(y)$ is continuous in this interval, is

$$s = \int_c^d \sqrt{1 + [g'(y)]^2} dy = \int_b^a \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy. \quad (6.2)$$



FIG. 15

The formula for arc length for a curve given by parametric equations can be derived from (6.1) or (6.2), and agrees with Eq. 7.1, Chapter 6. Let

$$x = f(t), \quad y = g(t),$$

and let the initial and terminal points of the arc correspond to t_1 and t_2 , respectively. It is convenient to denote differentiation with respect to the parameter t by a dot: $dy/dt = \dot{y}$, $dx/dt = \dot{x}$. Now,

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{\dot{y}}{\dot{x}}, \quad \text{and} \quad dx = \left(\frac{dx}{dt}\right) dt = \dot{x} dt.$$

Substituting in (6.1), we obtain

$$s = \int_{t_1}^{t_2} \sqrt{1 + \left(\frac{\dot{y}}{\dot{x}}\right)^2} \dot{x} dt = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2} dt. \quad (6.3)$$

Formula 6.1 can be used only with a curve that consists of a single-valued branch, i.e., one that is intersected just once by any ordinate drawn between a and b . A similar restriction applies to (6.2). Formula 6.3 is much less restricted and can be used for curves having vertical

tangents, loops, and other anomalies, as long as $f(t)$ and $g(t)$ have continuous derivatives. Figure 15 shows some types of curves whose lengths might be found by this formula, but to which (6.1) or (6.2) could be applied only after breaking the problem into several parts. Curves for which the integral (6.3) exists are said to be *rectifiable*. To *rectify* such a curve is to calculate its length.

The formulas for the length of an arc can also be derived from a sequence of approximations, each of which is an inscribed broken line, made up of chords instead of tangents, as shown in Fig. 16 (see Exercise 11).

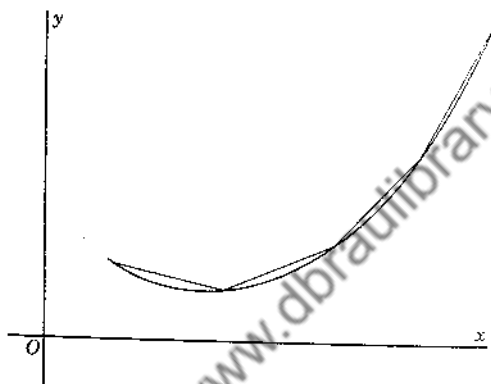


FIG. 16

EXAMPLE. Find the length of one arch of the cycloid

$$x = a(\theta - \sin \theta), \quad y = a(1 - \cos \theta).$$

SOLUTION. Differentiating,

$$dx/d\theta = \dot{x} = a(1 - \cos \theta), \quad dy/d\theta = \dot{y} = a \sin \theta.$$

Substituting in formula 6.3, we have

$$\begin{aligned} s &= \int_0^{2\pi} a \sqrt{(1 - \cos \theta)^2 + \sin^2 \theta} \, d\theta \\ &= a \int_0^{2\pi} \sqrt{2 - 2 \cos \theta} \, d\theta \\ &= a \int_0^{2\pi} \sqrt{2 \cdot 2 \sin^2 \frac{\theta}{2}} \, d\theta = 2a \int_0^{2\pi} \sin \frac{\theta}{2} \, d\theta \\ &= -4a \left[\cos \frac{\theta}{2} \right]_0^{2\pi} \\ &= 8a. \end{aligned}$$

EXERCISES

- Find the length of the circumference of a circle whose radius is r , by integration.
- Find the length of the curve $y^2 = x^3$ from $(1, -1)$ to $(4, -8)$.
- Find the length of the curve $x = t^3, y = 3t^2$, from $t = 1$ to $t = 3$.
- Find the length of the catenary $y = \frac{a}{2}(e^{x/a} + e^{-x/a})$ from $x = 0$ to $x = x_1$.
- Find the complete length of the cardioid $x = 2 \cos \theta - \cos 2\theta - 1$
 $y = 2 \sin \theta - \sin 2\theta$.
- Find the length of the curve $x = e^\theta \sin \theta$,
 $y = e^\theta \cos \theta$, from $(0, 1)$ to $(e^a \sin a, e^a \cos a)$.
- Find the complete length of the hypocycloid $x = a \cos^3 \theta, y = a \sin^3 \theta$.
- If from a spool 1 unit in radius a thread is unwound, which is kept constantly taut, its end describes the curve shown in Fig. 17. Parametric equations of this curve can be derived as follows.

$$x = ON \quad BP = \cos \theta + AP \sin \theta,$$

$$y = NA = BA = \sin \theta - AP \cos \theta.$$

Since AP is the length of thread unwound,
 $AP = \text{arc } AC = \theta$. Hence,

$$x = \cos \theta - \theta \sin \theta$$

$$y = \sin \theta - \theta \cos \theta.$$

This curve is called the *involute of the circle*.

(a) Show that the parametric equations given are correct if θ is a second quadrant angle.

(b) Find the length of the path of the end of the thread when one complete turn is unwound from the spool.

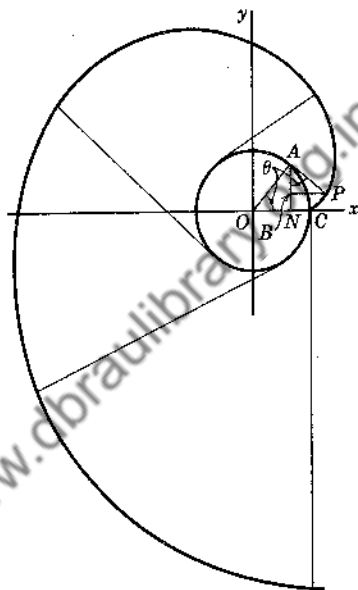


FIG. 17

- Find the length of the curve $y = \log(1 - x^2)$ from $x = 0$ to $x = .8$.
- Find the length of the curve $y = \log x$ from $x = 1$ to $x = e$.
- Derive formula 6.1 as follows, using Fig. 16.

(1) Express the length of a chord by $\sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{1 + (\Delta y/\Delta x)^2} \Delta x$.

(2) Use the mean value theorem to replace $\Delta y/\Delta x$ by $f'(x_i')$.

(3) Find $\lim_{n \rightarrow \infty} \sum_{i=1}^n \sqrt{1 + [f'(x_i')]^2} \Delta x$.

7. Arc Length as a Function. If we think of the arc length on a given curve as measured from some fixed initial point, the length is determined by the position of the terminal point. This corresponds to fixing the lower limit in the integral that gives the length, and letting the upper limit be variable. Thus, from (6.1),

$$s = \phi(\xi) = \int_a^\xi \sqrt{1 + [f'(x)]^2} dx. \quad (7.1)$$

It follows from the fundamental theorem (Art. 3, Chapter 8), that

$$ds/d\xi = \sqrt{1 + [f'(\xi)]^2}.$$

Since the particular letter that denotes a given variable is immaterial, we can write equally well,

$$Hence, \quad ds/dx = \sqrt{1 + [f'(x)]^2}. \quad (7.2)$$

$$Similarly, from (6.2), \quad ds = \sqrt{1 + [f'(x)]^2} dx. \quad (7.3)$$

$$and \quad ds/dy = \sqrt{1 + [g'(y)]^2}, \quad (7.4)$$

$$ds = \sqrt{1 + [g'(y)]^2} dy. \quad (7.5)$$

From the parametric formula, 6.3,

$$and \quad ds/dt = \sqrt{\dot{x}^2 + \dot{y}^2}, \quad (7.6)$$

$$ds = \sqrt{\dot{x}^2 + \dot{y}^2} dt. \quad (7.7)$$

8. Curvature. In Chapter 4 we found that the second derivative, d^2y/dx^2 , is a measure of the rate of turning of the tangent. This measure is doubly dependent upon the relationship of the curve with the coordinate system since (1) it is the rate with respect to the changing abscissa x , and (2) it is the rate of change of the slope, which is the tangent of the angle that the tangent to the curve makes with the x -axis. We now consider a measure of the turning of the tangent line that is completely independent of the coordinate system, depending only upon the curve itself. This measure is called the *curvature*.

Imagine a curve, with no particular coordinate system in evidence. As a point moves along the curve, the direction of its motion changes. In order to specify the direction of motion it is necessary to compare it with some fixed direction, e.g., due north, as when giving the course of a ship. The rate of change of direction is, therefore, the rate of change of an angle measured from the fixed direction. Now, although the size of the angle describing the given direction will depend upon the arbitrarily fixed direction, its rate of change is independent of the fixed direction, e.g., a course changes in the same way whether referred to due east or due north. We may therefore let the fixed direction be that of the positive x -axis in any rectangular coordinate system. The angle giving the direction of the curve at any point P is then the angle α between the tangent at P and the positive x -axis (Fig. 18). Either of two supplementary angles may be denoted by α , according to the sense of motion along the curve. We have to find the rate of change of α .

But with respect to what variable is this rate to be calculated? If the result is to be independent of the coordinate system, it cannot be with respect to x or y . Arc length, s , belongs intrinsically to the curve and is independent of the coordinates. Hence, we shall try to find $d\alpha/ds$. It is this quantity that is called the *curvature*, which we shall denote by K , writing

$$K = d\alpha/ds.$$

Of course, we know α only as $\tan^{-1} dy/dx$, and this quantity is closely tied to a coordinate system. Likewise, s is defined in terms of definite coordinates, e.g., by equation 7.1. Because of these facts, our system of coordinates will enter into the computation of K . The result, however, must be independent of the coordinates, since we have constructed it so that it would be. Consequently, if the same curve were to be represented in different ways on two different coordinate systems, the curvature K would be the same in both systems.

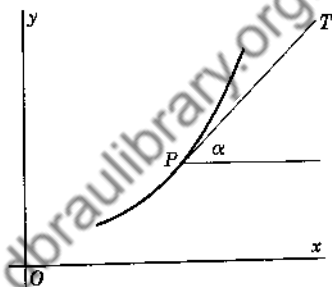


FIG. 18

In order to derive a formula for K in terms of x and y , we write

$$\begin{aligned} \frac{d\alpha}{ds} &= \frac{d\alpha}{dx} \frac{ds}{dx} = \left(\frac{d}{dx} \tan^{-1} y' \right) \frac{ds}{dx} \\ &= \frac{y''}{1 + (y')^2} \sqrt{1 + (y')^2}. \end{aligned}$$

Hence,

$$K = \frac{y''}{[1 + (y')^2]^{\frac{3}{2}}}. \quad (8.1)$$

The choice of sign in the denominator of this expression for K is arbitrary. Giving this radical a positive sign is equivalent to assigning a positive sign to ds/dx , i.e., assigning the positive sense on the curve to be that toward the positive x -direction. With this convention, the value of K found from (8.1) is positive if y'' is positive, in which case the direction of concavity is upward. Similarly, a negative result from (8.1) is associated with downward concavity.

A formula similar to (8.1) can be derived for the curvature when the curve is given by $x = g(y)$. It is

$$K = \frac{x''}{[1 + (x')^2]^{\frac{3}{2}}}. \quad (8.2)$$

where primes denote differentiation with respect to y . If the denominator in (8.2) is taken positively, the resulting value of K is positive if the curve is concave to the right and negative if it is concave to the left. The values of K from (8.1) and (8.2) can be made to agree by choosing the denominator in (8.2) positively if $x' < 0$ and negatively if $x' > 0$.

For a curve in parametric form: $x = f(t)$, $y = g(t)$, we again find $d\alpha/ds$, using $\alpha = \tan^{-1}(\dot{y}/\dot{x})$ and $ds/dt = \sqrt{\dot{x}^2 + \dot{y}^2}$. We have

$$\frac{d\alpha}{ds} = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{\dot{x}^2 + \dot{y}^2}$$

and, since

$$\frac{d\alpha}{ds} = \frac{d\alpha}{dt} \bigg/ \frac{ds}{dt},$$

$$K = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}. \quad (8.3)$$

As for Eq. 8.1, the choice of the positive sign of the denominator in (8.3) corresponds to the assumption that ds/dt is positive. If we were to change to another parameter, $u = -t$, the sign of K would change, since that of the numerator in (8.3) would change whereas that of the denominator would not (see Exercise 22, Art. 9). This would correspond to changing our point of view as to the positive direction along the curve.

In the general definition of curvature,

$$K = \frac{d\alpha}{ds} = \frac{d\alpha}{dt} \bigg/ \frac{ds}{dt}.$$

If we take ds/dt with the positive sign, K has the same sign as $d\alpha/dt$. Since the angle α is customarily measured positively in the counterclockwise direction, it follows that positive curvature corresponds to a bending of the curve in the counterclockwise direction, whereas negative curvature corresponds to a bending of the curve in the clockwise direction. If we reverse the positive direction along the curve, for example, by replacing t

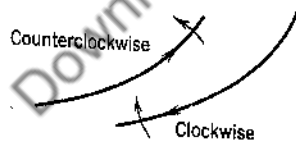


FIG. 19

by $-u$, the sign of K changes; at the same time, the direction of bending of the curve is reversed, changing from counterclockwise to clockwise, or vice versa, as shown in Fig. 19.

The unit of measure of K is radians per unit length. Thus, if distances are measured in inches, K is in radians per inch.

9. Radius and Circle of Curvature. Let us find the curvature of a circle of radius R . This circle can be represented parametrically by

$$x = R \cos t, \quad y = R \sin t.$$

Formula 8.3 then gives

$$K = \frac{(-R \sin t)(-R \sin t) - R \cos t(-R \cos t)}{(R^2 \sin^2 t + R^2 \cos^2 t)^{\frac{3}{2}}} \\ = 1/R.$$

In accordance with this result, we define the *radius of curvature* of a curve at a point to be the reciprocal of the absolute value of the curvature.

$$R = 1/|K|.$$

If, at a point of a curve, we construct a circle having the same tangent as the curve, with its radius $= R$, and its center on the concave side of the curve, we obtain a circle that agrees with the given curve at this point both in direction and in curvature. This circle is called the *circle of curvature* of the curve at the point, and its center is called the *center of curvature* of the curve for that point.

EXAMPLE 1. Find the curvature of the cycloid

$$x = a(\theta - \sin \theta),$$

$$y = a(1 - \cos \theta),$$

at any point.

SOLUTION. We have

$$dx/d\theta = a(1 - \cos \theta), \quad dy/d\theta = a \sin \theta.$$

$$d^2x/d\theta^2 = a \sin \theta, \quad d^2y/d\theta^2 = a \cos \theta.$$

Substituting these quantities in formula 8.3,

$$K = \frac{a(1 - \cos \theta)a \cos \theta - a \sin \theta \cdot a \sin \theta}{[a^2(1 - \cos \theta)^2 + a^2 \sin^2 \theta]^{\frac{3}{2}}} \\ = -\frac{1}{2\sqrt{2} a \sqrt{1 - \cos \theta}} = \frac{1}{4a |\sin \theta/2|}.$$

For the radius of curvature

$$R = 1/|K| = |4a \sin \theta/2|.$$

The radius of curvature is clearly a maximum at $\theta = \pi$, the topmost point of the cycloid. At $\theta = 0, 2\pi$, etc., R becomes 0. At these points, the cycloid has its cusps.

EXAMPLE 2. Find the radius of curvature of the curve $y = x^4 - 2x^2$ at $(0, 0)$, $(1, -1)$, and $(\sqrt{2}, 0)$.

SOLUTION. Differentiating,

$$y' = 4x^3 - 4x = 4x(x^2 - 1),$$

$$y'' = 12x^2 - 4 = 4(3x^2 - 1).$$

At $(0, 0)$: $y' = 0$, $y'' = -4$, and hence $K = -4/1$ and $R = 1/4$.

At $(1, -1)$: $y' = 0$, $y'' = 8$, and hence $K = 8/1$ and $R = 1/8$.

At $(\sqrt{2}, 0)$: $y' = 4\sqrt{2}$, $y'' = 20$, and hence $K = 20/(1 + 32)^{3/2} = 20/33^{3/2}$ and $R = 33^{3/2}/20 = 9.4$, approximately.

Conceiving the positive description of the curve to be from left to right, negative curvature corresponds to downward concavity, as at $(0, 0)$, and

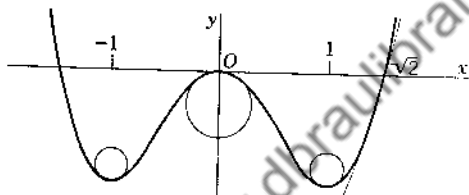


FIG. 20

positive curvature to upward concavity, as at $(1, -1)$. Figure 20 shows the curve with the circles of curvature at $(0, 0)$ and $(1, -1)$ and an arc of the circle of curvature at $(\sqrt{2}, 0)$.

EXERCISES

Find K and R for each of the curves in Exercises 1–12 at the points specified.

- $x = 3t$, $y = t^2 - 1$; $t = 1$, $t = 2$.
- $y = x^2$; $(0, 0)$, $(1, 1)$, $(2, 4)$.
- $y = x^3 - 12x$; $(0, 0)$, $(2, -16)$.
- $xy = 16$; $(4, 4)$, $(2, 8)$.
- $x = 4 \cos^3 t$, $y = 4 \sin^3 t$; $t = \pi/4$.
- $y = \log x$; $(1, 0)$, $(1/\sqrt{2}, -\frac{1}{2} \log 2)$.
- $y = \sin x$; $(\pi/2, 1)$, $(\pi, 0)$.
- $y = \frac{1}{2}(e^x + e^{-x})$; $(0, 1)$.
- $y = \log \sec x$; $(0, 0)$, $(\pi/3, 2)$.
- $x^2/a^2 + y^2/b^2 = 1$; $(0, b)$, $(a, 0)$, $(a/\sqrt{2}, b/\sqrt{2})$.
- $x = 2 \cos \theta - \cos 2\theta - 1$, $y = 2 \sin \theta - \sin 2\theta$. $\theta = \pi/3$, $\theta = \pi$.
- $x = 3t/(1 + t^3)$, $y = 3t^2/(1 + t^3)$. $t = 0$, $t = 1$.
- The curvature of a straight line is, of course, always 0. Show this by applying one of the formulas for K to a linear function.

14. The curvature of a circle is, of course, constant. Show this by use of the equation $x^2 + y^2 = r^2$.

15. How does the curvature of a circle of radius 4 compare with that of a circle of radius 1?

16. A circle has its curvature equal to 4 rad per ft. How large is the circle?

17. Show that the curvature is 0 at a point of inflection if the curvature is continuous. Is every point of 0 curvature a point of inflection?

18. Find the point at which the curve $y = e^x$ turns most rapidly, and find the radius of curvature at this point.

19. Find the points where the curve $y = \sin x$ turns most rapidly and those where it turns least rapidly.

20. Find the critical points of the curve $y = x^3 - 3x$, and find the curvature at each of these points. Show that the curvature does not have its maximum value at a critical point.

21. Find the curvature of the involute of the circle, the equations of which were derived in Exercise 8, Art. 6, at any point $\theta = \theta_1$. Draw the circle of curvature at the point where its radius is 1. What can be said about the circle of curvature at the point P , as P moves toward C as a limiting position (Fig. 17)?

22. Show that the formula for K is unchanged by changing the parameter. That is, show that if $t = \phi(u)$

$$\frac{f'(t)g''(t) - f''(t)g'(t)}{[f'(t)^2 + g'(t)^2]^{\frac{3}{2}}} = \frac{F'(u)G''(u) - G'(u)F''(u)}{[F'(u)^2 + G'(u)^2]^{\frac{3}{2}}},$$

where $F(u) = f[\phi(u)]$, $G(u) = g[\phi(u)]$.

23. If the transformation in Exercise 22 is $t = -u$, show that each first derivative changes sign, whereas the second derivatives do not. Hence, show that if the denominator of K is taken with a positive sign, the sign of K changes as a result of this transformation.

24. Prove the statement in the text about the sign of K in (8.2).

10. The Area Problem Reviewed. The most natural geometrical application of the definite integral is to the problem of area. This application has been studied at several points in earlier chapters. In this article we shall review it briefly and show how to find areas bounded by curves expressed parametrically.

For the area bounded by a curve $y = f(x)$, the x -axis, and two ordinates $x = a$ and $x = b$, the formula is

$$A = \int_a^b y \, dx = \int_a^b f(x) \, dx. \quad (10.1)$$

It was shown in Chapter 9 how this formula could be adapted to find areas below the x -axis, as well as those above, and areas that are bounded both above and below by curves. The adaptation to areas bounded by a curve $x = g(y)$, the y -axis, and two abscissas $y = c$ and $y = d$ was also shown. In this case the formula is

$$A = \int_c^d x \, dy = \int_c^d g(y) \, dy.$$

Now, let us suppose a curve is expressed parametrically by $x = \phi(t)$, $y = \psi(t)$. The area, illustrated in Fig. 21, can be found by formula 10.1,

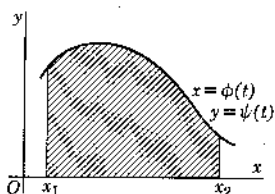


FIG. 21

$$A = \int_{x_1}^{x_2} y \, dx,$$

where y is replaced by $\psi(t)$ and dx by $\phi'(t) \, dt$. That is,

$$A = \int_{t_1}^{t_2} \psi(t) \phi'(t) \, dt, \quad (10.2)$$

in which t_1 and t_2 are the values of t that correspond to x_1 and x_2 , respectively.

EXAMPLE. Find the area of a quadrant of the ellipse

$$x = a \cos t, \quad y = b \sin t.$$

SOLUTION. From Fig. 22 we have to find

$$\begin{aligned} A &= \int_0^a y \, dx = \int_{\pi/2}^0 b \sin t \frac{d}{dt} (a \cos t) \, dt \\ &= \int_{\pi/2}^0 ab (-\sin^2 t) \, dt. \end{aligned}$$

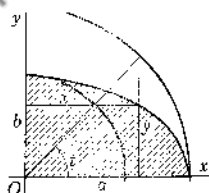


FIG. 22

The limits are obtained by noting that $x = 0$ when $a \cos t = 0$, i.e., $t = \pi/2$; and $x = a$ when $a \cos t = a$, i.e., $t = 0$. Changing the order of the limits and the sign of the integrand, in accordance with Theorem 2, Art. 10, Chapter 7,

$$A = ab \int_0^{\pi/2} \sin^2 t \, dt = \frac{\pi ab}{4}.$$

An important extension of formula 10.2 enables us to find the area within a closed curve by a single integration. Suppose the closed curve in Fig. 23 is expressed by

$$x = \phi(t), \quad y = \psi(t).$$

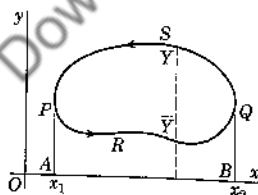


FIG. 23

Let $t = t_1$ correspond to the point P at which the abscissa has its least value x_1 , and $t = t_2$ to the point Q where the abscissa has its greatest value x_2 . Since the curve is closed, there must be an interval of values of t , say t_1 to t_2 , such that, as t varies from t_1 to t_2 , the whole curve is described exactly once. Consequently, $\phi(t_2) = \phi(t_1)$ and $\psi(t_2) = \psi(t_1)$. For explicitness we shall

suppose, also, that the curve is described counterclockwise, i.e., so that the enclosed area is always to the left as one traverses the boundary.

The area enclosed is

$$A = \text{Area } PABQSP - \text{Area } PABQRP = \int_{x_1}^{x_2} Y dx - \int_{x_1}^{x_2} \bar{Y} dx,$$

where Y denotes values of y on the arc PSQ and $\bar{Y}$ the values of y on the arc PRQ . Replacing these integrals by their parametric forms, we have

$$\begin{aligned} A &= \int_{t_1}^{t_2} \psi(t)\phi'(t) dt - \int_{t_1}^{t_2} \bar{\psi}(t)\bar{\phi}'(t) dt \\ &= - \int_{t_1}^{t_2} \psi(t)\phi'(t) dt - \int_{t_1}^{t_2} \bar{\psi}(t)\bar{\phi}'(t) dt, \end{aligned}$$

by Theorem 2, Art. 10, Chapter 7. This becomes, by Theorem 3, Art. 10, Chapter 7,

$$A = - \int_{t_1}^{t_2} \psi(t)\phi'(t) dt \quad (10.3)$$

or

$$A = - \int_{t_1}^{t_2} y dx, \quad (10.4)$$

in which it is understood that y and dx must be replaced by their equals in terms of t before integration.

Since this formula was derived on the assumption of counterclockwise description of the curve, if the description is in the opposite direction the sign of the integral is reversed. In the application of the formula, the interval of values of t may be so chosen as to correspond to any initial point whatever, with t varying so as to describe the complete curve just once. It is assumed, further, that the boundary curve does not intersect itself.

Applying formula 10.3, we can find the area of the entire ellipse in the illustrative example. We have

$$\begin{aligned} A &= - \int_0^{2\pi} b \sin t (-a \sin t) dt \\ &= ab \int_0^{2\pi} \sin^2 t dt = \pi ab. \end{aligned}$$

EXERCISES

1. Find the area bounded by the curve $x^{\frac{1}{2}} + y^{\frac{1}{2}} = a^{\frac{1}{2}}$ and the coordinate axes.
2. Find the area bounded by the catenary $y = (e^{x/a} + e^{-x/a})$, the x -axis, and the ordinates $x = -a$ and $x = a$.

3. Find the area bounded by the curve $x = \log t$, $y = 1/t$, the coordinate axes, and the ordinate $x = a$. If we take a larger and larger, how is the area affected?
4. Find the area bounded by one arch of the cycloid $x = a(\theta - \sin \theta)$, $y = a(1 - \cos \theta)$, and the x -axis.
5. Find the area bounded by one arch of the curve $x = a\theta$, $y = a(1 - \cos \theta)$, and the x -axis.
6. Find the area in the second quadrant bounded by the cardioid, $x = 2 \cos \theta - \cos 2\theta - 1$, $y = 2 \sin \theta - \sin 2\theta$, and the coordinate axes.
7. Find the area in the first quadrant bounded by the cardioid in Exercise 2 and the y -axis. Hence, using the result of Exercise 2, find the complete area within the cardioid. [Hint: $A = \int_0^{\pi/2} x(dy/d\theta) d\theta$.]
8. Find the complete area within the cardioid of Exercise 2 by means of formula 10.3.

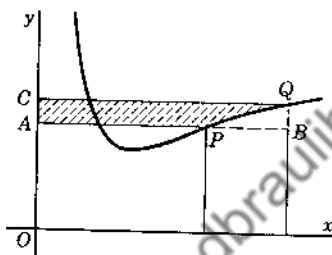


FIG. 24

9. Find the area bounded by the hypocycloid $x = a \cos^3 \theta$, $y = a \sin^3 \theta$.
10. Find the area of the loop of the curve $y^2 = x^2(a + x)/(a - x)$.
11. Find the area bounded by the curve $y = \log x + 1/x$, the y -axis, and two abscissas whose lengths are 2 and 3, respectively (Fig. 24). [Hint: Find the area $ABQC$ and PBQ and subtract.]
12. Starting with the formula $A = \int_{y_1}^{y_2} x dy$, and taking steps similar to those by which formula 10.3 was derived, obtain the formula

$$A = \int_{t_1}^{t_2} \phi(t)\psi'(t) dt \quad (10.5)$$

$$= \int_{t_1}^{t_2} x dy,$$

provided that the complete boundary is described just once as t varies from t_1 to t_2 .

11. Volumes of Revolution. Shell Method. Volumes of revolution, like areas, have been studied in earlier chapters. We found that the volume generated by revolving about the x -axis an area bounded by the curve $y = f(x)$, the x -axis, and two ordinates $x = x_1$ and $x = x_2$, could be expressed by the integral,

$$V = \int_{x_1}^{x_2} \pi y^2 dx = \int_{x_1}^{x_2} \pi [f(x)]^2 dx. \quad (11.1)$$

We obtained this formula by thinking of the volume as the limit of the sum of approximating cylindrical discs, each having a volume of $\pi[f(x_i)]^2 \Delta x$.

The transformation of formula 11.1 for a curve which is expressed parametrically is similar to that shown in Art. 10. If the curve is expressed by $x = \phi(t)$, $y = \psi(t)$,

$$V = \int_{x_1}^{x_2} \pi y^2 dx = \int_{t_1}^{t_2} \pi[\psi(t)]^2 \phi'(t) dt, \quad (11.2)$$

where t_1 and t_2 are the values of t that correspond to $x = x_1$ and $x = x_2$, respectively.

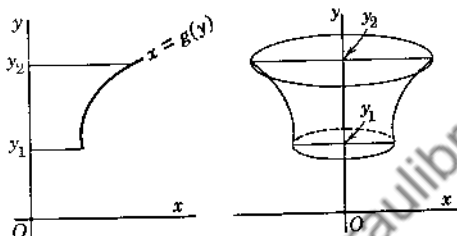


FIG. 25

If the area shown in Fig. 25 is revolved about the y -axis, the volume is given by

$$V = \int_{y_1}^{y_2} \pi x^2 dy = \int_{y_1}^{y_2} \pi[g(y)]^2 dy. \quad (11.3)$$

If this curve were given parametrically by $x = \phi(t)$, $y = \psi(t)$, the volume would be found by

$$V = \int_{t_1}^{t_2} \pi[\phi(t)]^2 \psi'(t) dt, \quad (11.4)$$

where t_1 and t_2 correspond to $y = y_1$ and $y = y_2$, respectively.

A different method of finding a volume of revolution is sometimes more convenient than that which makes use of the above formulas. We shall consider it first in an extremely simple example, namely, that of calculating the volume of a cone.

EXAMPLE 1. Calculate the volume of the cone generated by revolving the triangle formed by the coordinate axes and the line $2x + y = 6$ about the y -axis (Fig. 26).

SOLUTION. Instead of thinking of this volume as approximated by a sum of cylindrical plates, we approximate it by a sum of cylindrical shells. Let the inner radius of a shell be $\underline{x}_i$, the outer radius $\bar{x}_i$, and the thickness Δx_i . If we let the height of the shell be that of a point of the hypotenuse of the triangle for some value $x = x_i$, between $\underline{x}_i$ and $\bar{x}_i$, this height is

$y_i = 6 - 2x_i$, and the volume of the shell is found by subtracting the volumes of two cylinders with height y_i , one with radius $\underline{x}_i$ and the other with radius $\bar{x}_i$. Denoting this volume by $\Delta_i V$, we have

$$\begin{aligned}\Delta_i V &= \pi \bar{x}_i^2 y_i - \pi \underline{x}_i^2 y_i = \pi y_i (\bar{x}_i + \underline{x}_i)(\bar{x}_i - \underline{x}_i) \\ &= \pi(6 - 2x_i)2x_i' \Delta_i x_i\end{aligned}$$

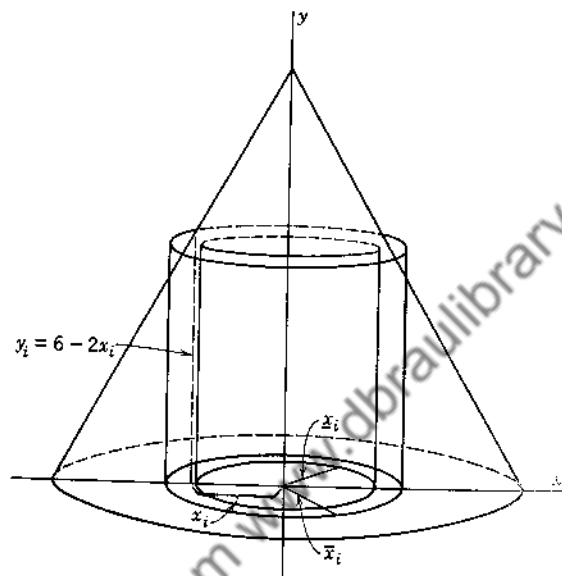


FIG. 26

where $x_i' =$ the average of $\underline{x}_i$ and $\bar{x}_i$, and $\Delta_i x =$ the thickness, $\bar{x}_i - \underline{x}_i$. Here x_i and x_i' may not be the same, although both lie in the interval $\Delta_i x$. We now have the approximate value of the complete volume

$$V = \sum_{i=1}^n 2\pi(6 - 2x_i)x_i' \Delta_i x.$$

This sum has the form that leads to the integral $\int_0^3 2\pi(6 - 2x)x \, dx$. We thus conclude that

$$V = \int_0^3 2\pi(6 - 2x)x \, dx = 18\pi.$$

The limits 0 and 3 are the least and greatest x , respectively, for any shells in the figure.

The two methods of finding volumes of revolution, the disc method and the shell method, lead to two definitions of such volumes by the limits

of two different sums. These definitions, however, can be shown to be equivalent.

The shell method leads to two new formulas for volumes of revolution.

(1) If the area is revolved about the x -axis,

$$V = 2\pi \int_{y_1}^{y_2} h y \, dy. \quad (11.5)$$

(2) If the area is revolved about the y -axis,

$$V = 2\pi \int_{x_1}^{x_2} h x \, dx. \quad (11.6)$$

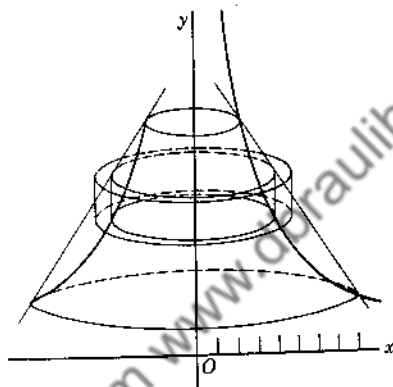


FIG. 27

In these formulas, h = height of the shell. This is determined from the boundary of the area revolved. The limits are the numerical values of the least and greatest radii, respectively.

Formulas 11.5 and 11.6 are transformed for use with parametric equations in the same way as were formulas 11.1 and 11.3.

EXAMPLE 2. Find the volume generated by revolving about the y -axis the area bounded by the hyperbola $xy = 16$ and the line $x + y = 10$.

SOLUTION. We use (11.6) with h equal to the difference between the height of the line and that of the hyperbola. Since the height of the line is $10 - x$, and that of the hyperbola is $16/x$, we have $h = 10 - x - 16/x$. The limits of integration are found by solving the equations $x + y = 10$ and $xy = 16$ simultaneously. This gives us the points of intersection, $(2, 8)$ and $(8, 2)$. An elemental shell is shown in Fig. 27. The volume is

$$V = 2\pi \int_2^8 x(10 - x - 16/x) \, dx = 72\pi.$$

EXERCISES

Find the volume generated by revolving each of the areas in Exercises 1-14 about the axis indicated.

1. The area bounded by $y = x^2$, the x -axis, and $x = 1$, about the y -axis.
2. The area bounded by the parabola $y = x^2/4$, the x -axis, and the line $x = 2$, about the y -axis.
3. The area bounded by the parabola $y = 8 - x^2$ and the line $y = 6$, about the x -axis.
4. The area bounded by the curve $y = e^x$, the coordinate axes, and the line $x = -1$, about the y -axis.
5. The area bounded by one arch of the curve $y = \sin x$ and the x -axis, about the y -axis.
6. The area bounded by $y = \log x$, the coordinate axes, and the line $y = 1$, about the x -axis.
7. The area $ABCD$ in Fig. 28, about the y -axis.

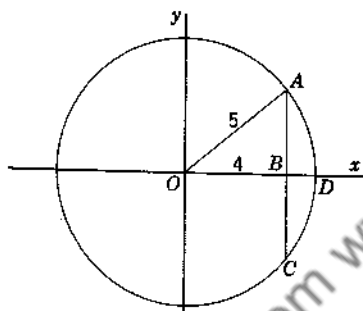


FIG. 28

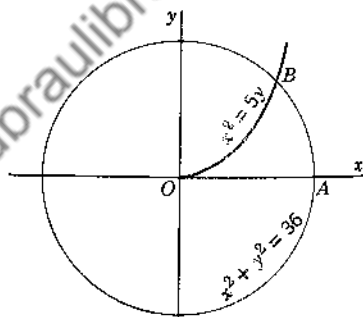


FIG. 29

8. The area bounded by the two parabolas $y = 4 - x^2$ and $2y = x^2 - 4$, about the y -axis.
9. The area OAB in Fig. 29, about the x -axis.
10. The area bounded by one arch of the cycloid $x = a(\theta - \sin \theta)$, $y = a(1 - \cos \theta)$, about the y -axis.
11. The area of the circle $(x - a)^2 + y^2 = r^2$, $a > r$, about the y -axis.
12. The area bounded by the curve $y = e^{-x^2}$, the coordinate axes, and the line $x = 1$, about the y -axis.
13. The area bounded by the parabolas $y^2 = x$ and $x^2 = y$, about the y -axis.
14. The sector of a circle of radius r having a central angle of 45° , about a bounding radius.
15. Evaluate the volume in Exercise 12 by the disc method.

12. Surfaces of Revolution. Suppose a curve, which we shall assume lies entirely on one side of the x -axis, revolves about the x -axis. The curve (not the area bounded by it) sweeps out, or generates, a surface of

revolution. We wish to calculate the area of this surface. Let us consider an example.

EXAMPLE 1. Find the area of the surface of revolution generated by revolving about the x -axis the arc of the hyperbola $xy = 1$ that lies between $(1, 1)$ and $(2, \frac{1}{2})$.

SOLUTION. We have no formula for the area of a curved surface like this, nor, indeed, a definition of it. We shall attempt to reach a definition satisfactory to geometric intuition, as we have previously done for plane areas, lengths of arcs, and volumes of revolution, by means of the limit of a sequence of approximating sums. We think of the surface as

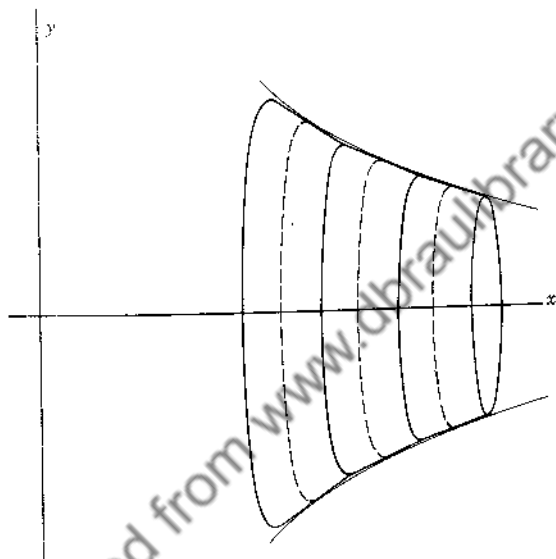


FIG. 30

approximated by a series of bands, each in the form of a frustum of a conical surface tangent to the surface along a circle halfway between the bases (Fig. 30). (The requirement that the circle of tangency be halfway between the bases is a convenience rather than a necessity.) From the elementary formula for the area of a frustum we can then express the sum of the area of the approximating bands. This series of bands can be made to approximate the surface more and more closely by taking a larger and larger number of bands and at the same time decreasing their widths. We shall take the area of the curved surface to be the limit of the sequence of sums of areas of the bands, which will be in the form of an integral.

In order to set up the sum from which the integral is to be obtained,

let us examine one of the approximating bands. The lateral area of the frustum in Fig. 31, denoted by $\Delta_i A$, is, by a formula from solid geometry,

$$\Delta_i A = 2\pi l_i h_i,$$

in which h_i is the height of the midpoint of PQ , i.e., of the point of tangency, and l_i is the length of PQ . Since the point of tangency lies on the curve $y = 1/x$, its height is

$$h = 1/x'_i, \text{ wherein } x'_i = (x_{i-1} + x_i)/2.$$

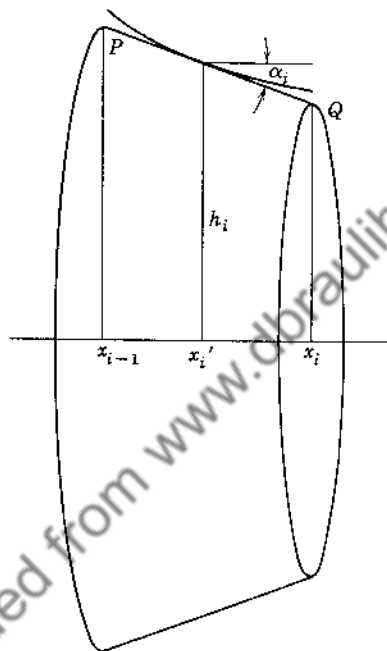


FIG. 31

Moreover, if α_i denotes the angle of inclination of the tangent,

$$l_i = \sec \alpha_i \cdot \Delta_i x.$$

Since

$$\tan \alpha_i = (dy/dx)_{(x=x'_i)} = -1/(x'_i)^2,$$

$$\sec \alpha_i = \sqrt{1 + 1/(x'_i)^4}.$$

Hence, the area of the i th approximating band is

$$\Delta_i A = 2\pi \sqrt{1 + 1/(x'_i)^4} \cdot (1/x'_i) \Delta_i x.$$

Defining the area of the surface to be the limit of the sum of the areas of all approximating bands as their number tends to infinity, we have

$$A = 2\pi \int_1^2 \sqrt{1 + 1/x^4} \cdot (1/x) dx$$

$$= \pi \int_1^4 \frac{\sqrt{u^2 + 1}}{u^2} du, \quad \text{wherein } u = x^2.$$

Integrating by parts,

$$\int \frac{\sqrt{u^2 + 1}}{u^2} du = -\frac{1}{u} \sqrt{u^2 + 1} + \log(u + \sqrt{u^2 + 1}).$$

Hence,

$$A = \pi \left[\log(u + \sqrt{u^2 + 1}) - \frac{\sqrt{u^2 + 1}}{u} \right]_1^4$$

$$= 5.02, \text{ approximately.}$$

In order to generalize this example, we imagine a surface generated by revolving the section of the curve $y = f(x)$ between x_1 and x_2 about the x -axis. We shall assume that $x_1 < x_2$ and $f(x) > 0$ throughout the interval. Just as in the example, we obtain

$$A = \int_{x_1}^{x_2} 2\pi \sqrt{1 + [f'(x)]^2} f(x) dx, \quad (12.1)$$

or more briefly,

$$A = \int_{x_1}^{x_2} 2\pi y ds, \quad (12.2)$$

in which y and ds are to be replaced by their equals in terms of x . Formula 12.1 is both a definition of the area and a formula for calculating it.

For a curve having the equation $x = g(y)$, where $g(y) > 0$, rotated about the y -axis, formula 12.1 is replaced by

$$A = \int_{y_1}^{y_2} 2\pi \sqrt{1 + [g'(y)]^2} g(y) dy, \quad (12.3)$$

and formula 12.2 by

$$A = \int_{y_1}^{y_2} 2\pi x ds. \quad (12.4)$$

In these formulas we assume that $y_1 < y_2$.

For curves expressed parametrically by $x = \phi(t)$, $y = \psi(t)$, formulas 12.2 and 12.4 become

$$A = \int_{t_1}^{t_2} 2\pi y \sqrt{\dot{x}^2 + \dot{y}^2} dt \quad (12.5)$$

and

$$A = \int_{t_1}^{t_2} 2\pi x \sqrt{\dot{x}^2 + \dot{y}^2} dt, \quad (12.6)$$

respectively. In (12.5) we assume $y > 0$ and in (12.6), $x > 0$, and in both, $t_1 < t_2$. In these formulas, of course, we must substitute $x = \phi(t)$, $y = \psi(t)$, $\dot{x} = \phi'(t)$, $\dot{y} = \psi'(t)$, before integrating.

EXAMPLE 2. Find the area of the surface generated by revolving about the y -axis the part of the hypocycloid $x = a \cos^3 \theta$, $y = a \sin^3 \theta$, that lies in the first quadrant.

SOLUTION. Using formula 12.6, with $\dot{x} = -3a \cos^2 \theta \sin \theta$, $\dot{y} = 3a \sin^2 \theta \cos \theta$, we have

$$\begin{aligned} A &= \int_0^{\pi/2} 2\pi \cdot 3a \sqrt{\cos^4 \theta \sin^2 \theta + \sin^4 \theta \cos^2 \theta} \cdot a \cos^3 \theta d\theta \\ &= 6\pi a^2 \int_0^{\pi/2} \cos^4 \theta \sin \theta d\theta = \frac{6\pi a^2}{5} \end{aligned}$$

EXERCISES

1. Find, by integration, the area of the surface of the sphere generated by revolving the unit circle about its diameter.
2. A zone of a sphere is generated by revolving an arc of a circle about a diameter. Show that the area of a zone of height h on a sphere of radius r is the same for all such zones, and find this area. [Hint: Use the equation $x^2 + y^2 = r^2$ and integrate from x_1 to $x_1 + h$.]
3. Find the area of the surface generated by revolving about the x -axis the arc of the parabola $x = y^2$ between the points $(0, 0)$ and $(4, 2)$.
4. Find the area of the surface generated by revolving about the x -axis the arc of the curve $y = x^3$ between the points $(0, 0)$ and $(2, 8)$.
5. Find the area in Exercise 4 if the arc is revolved about the y -axis.
6. Find the area generated by revolving about the x -axis the arc of the catenary $y = \frac{1}{2}(e^x + e^{-x})$ between $x = -1$ and $x = 1$.
7. Find the area of the surface generated by revolving about the x -axis the arc of the curve $y = e^{-x}$ between $x = 0$ and $x = \frac{1}{2} \log 24$.
8. Find the area of the surface generated by revolving about the x -axis one arch of the curve $y = \sin x$.
9. Find the area of the surface generated by revolving one arch of the cycloid $x = a(\theta - \sin \theta)$, $y = a(1 - \cos \theta)$ about the x -axis.
10. Find the area of the ellipsoid of revolution generated by revolving an ellipse about (a) the major axis; (b) the minor axis.

15

Geometrical Applications Polar Coordinates

1. Introduction. Rectangular coordinate systems are used so frequently and with so wide a variety of geometrical and physical applications that one may unconsciously come to identify "coordinate system" with "rectangular coordinate system." Indeed, we have described the two fundamental problems of calculus in geometrical terms closely associated with rectangular coordinates. The tangent problem is directly related to the angle which the tangent to a curve makes with the x -axis. The area problem, as we have studied it so far, always reduces to the area between a curve and one of the rectangular coordinate axes.

This dependence upon rectangular coordinates is by no means essential to the development of calculus. In fact, the whole subject could be built up without any reference to geometrical elements. The notion of the derivative as the slope of the tangent and that of the integral as the area between a curve and one of the coordinate axes are merely interpretations of these calculus concepts. In another kind of coordinate system, different geometrical interpretations are not only possible but are demanded by the structure of the system. In the present chapter we shall consider the interpretations of some of the quantities of calculus in a *polar coordinate system*.

2. The Polar Coordinate System. The position of a point is determined in a polar coordinate system by means of its distance from a fixed point and its direction from this fixed point as compared with a fixed direction. Thus, in Fig. 1, the point P is 3 units from the fixed point, or *pole*, O , and the direction of OP is $\pi/6$ radians measured positively from the fixed direction OX . There is no reason why the line OX must extend horizontally to the right, and any other direction would serve equally well as the fixed direction.

The polar coordinates of a point are called, in general, r and θ , as indicated in Fig. 2. The coordinates of P in Fig. 1 are $(3, \pi/6)$. If r and θ are connected by some relation, $r = f(\theta)$, a locus is determined—the locus of all points satisfying this relation. This is usually a curve,

often called a *polar curve*, since it is drawn with reference to the fixed point and fixed direction of a polar coordinate system.

The functional relation, $r = f(\theta)$, differs in no respect from the functional relation, $y = f(x)$. It is only the graphical interpretations that

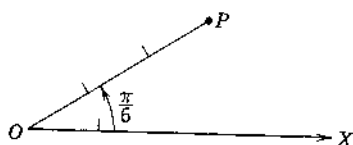


FIG. 1

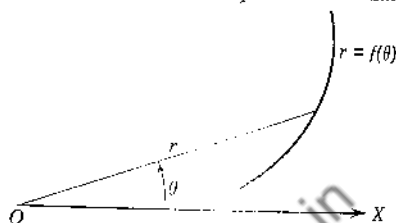


FIG. 2

differ if we choose to represent $r = f(\theta)$ on a polar system and $y = f(x)$ on a rectangular system. Actually, there is no reason why we could not interpret r and θ as rectangular coordinates, or x and y as polar coordinates, if we chose to do so. If $f(x)$ has a derivative, so has $f(\theta)$; and if $x_1 = \theta_1$, $f'(x_1) = f'(\theta_1)$. If the integral $\int_a^b f(x) dx$ exists, so does the integral $\int_a^b f(\theta) d\theta$, and $\int_a^b f(x) dx = \int_a^b f(\theta) d\theta$. If, however, r and θ are polar coordinates, $f'(\theta_1)$ does not represent the slope of the curve $r = f(\theta)$ at $\theta = \theta_1$, and $\int_a^b f(\theta) d\theta$ does not represent the area under the curve between $\theta = a$ and $\theta = b$.

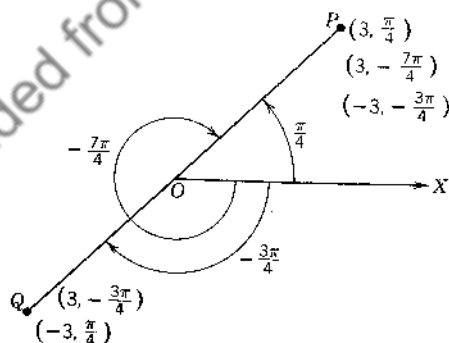


FIG. 3

In the present chapter we shall not further develop the subject of calculus by differentiating or integrating any new functions or introducing new concepts of calculus. Our whole task is that of geometrical interpretation of the concepts and results of calculus in a polar coordinate

system. We state here, for reference and review, some of the fundamental conventions regarding a polar coordinate system.

1. The positive description of the angle θ is counterclockwise.
2. The positive direction of r , the *radius vector*, is that of the terminal ray of the angle θ . The negative direction is, of course, the opposite of this, or differing from it by π radians.
3. Each point has infinitely many representations. Some different representations of the same point are indicated in Fig. 3.

4. If a rectangular coordinate system is superposed on a polar system, as shown in Fig. 4, a point P has the polar coordinates (r, θ) and the rectangular coordinates (x, y) which are related by the equations:

$$x = r \cos \theta, \quad y = r \sin \theta,$$

$$r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}(y/x).$$

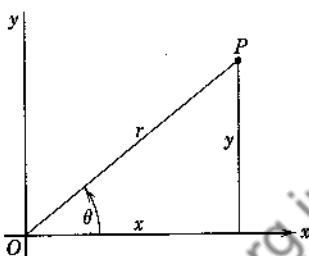


FIG. 4

3. The Tangent Problem in Polar Coordinates. The tangent problem is that of finding the direction of the tangent to a curve at a given point. When the curve was the graphical representation in rectangular coordinates of an equation $y = F(x)$, it was convenient to express this direction by means of the slope, or tangent of the angle which the tangent line made with the x -axis. If the curve is the graphical representation in polar coordinates of an equation $r = f(\theta)$, it will be more convenient to express this direction in terms of a different angle, namely angle ψ which the tangent line makes with the radius vector to the point of tangency (Fig. 5). After we have found this angle ψ it is a simple matter to find the angle α which the tangent makes with the fixed direction OX , since

$$\alpha = \psi + \theta - n\pi,$$

FIG. 5

for some integral value of n .

For the purpose of finding ψ we start, as in rectangular coordinates, with two points $P(r, \theta)$ and $Q(r + \Delta r, \theta + \Delta \theta)$ (Fig. 6). The direction of PQ is, of course, only a poor approximation to that of the tangent PT at P . Nevertheless, as Q approaches P , the direction of PQ approximates

that of PT more and more closely, and if we define the direction of PT to be the limiting direction approached by the direction of PQ as Q approaches P , this is no new definition. It is precisely the same, geometrically, as that by which the direction of the tangent was defined in Chapter 2. For clarity, it is necessary that this concept be couched in numerical terms related to the polar coordinate system. Accordingly, we define the numerical quantity, $\tan \psi$. As Q approaches P , the angle ϕ clearly approaches the angle intuitively accepted as ψ . Hence, we define $\tan \psi$ (and ψ) as follows.*

$$\tan \psi = \lim_{\Delta\theta \rightarrow 0} \tan \phi.$$

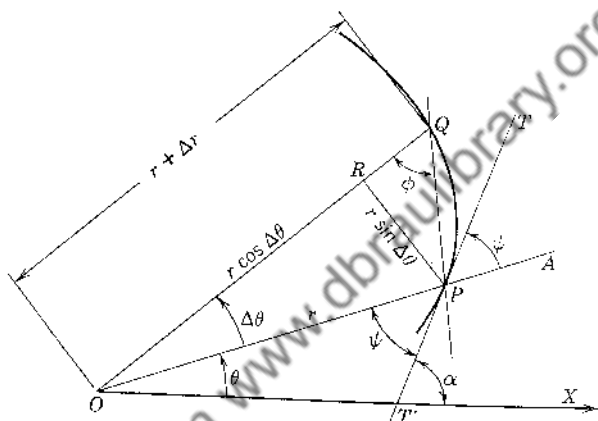


FIG. 6

In order to express $\tan \phi$ in terms of r and θ we draw the line PR perpendicular to OQ . Now,

$$OP = r, \quad OQ = r + \Delta r,$$

$$PR = r \sin \Delta\theta, \quad OR = r \cos \Delta\theta, \quad RQ = r + \Delta r - r \cos \Delta\theta.$$

Referring to the right triangle PQR ,

$$\tan \phi = \frac{r \sin \Delta\theta}{r + \Delta r - r \cos \Delta\theta} = \frac{r \sin \Delta\theta}{r(1 - \cos \Delta\theta) + \Delta r}.$$

* This is a new definition of an angle, giving the direction of the tangent line, independent of the former definition of the slope of the tangent. We have as much right to make one of these definitions as the other, but of course it is desirable that they be consistent, i.e., determine the same tangent line. It is not difficult to prove that this is so, although we shall not do so.

Dividing numerator and denominator by $\Delta\theta$,

$$\tan \phi = \frac{r \sin \Delta\theta / \Delta\theta}{r(1 - \cos \Delta\theta) / \Delta\theta + \Delta r / \Delta\theta}$$

As $\theta \rightarrow 0$, $\frac{\sin \Delta\theta}{\Delta\theta} \rightarrow 1$, and $\frac{1 - \cos \Delta\theta}{\Delta\theta} \rightarrow 0$, from the results in Art. 6, Chapter 10. At the same time, $\Delta r / \Delta\theta \rightarrow dr/d\theta = f'(\theta)$, which can be obtained from the equation of the curve. Hence,

$$\lim_{\Delta\theta \rightarrow 0} \tan \phi = \tan \psi = \frac{r}{dr/d\theta} = \frac{r}{f'(\theta)}$$

This result can also be written $\tan \psi = r(d\theta/dr)$, or in terms of the differentials $d\theta$ and dr , $\tan \psi = (r d\theta)/dr$.

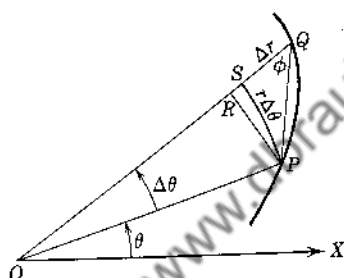


FIG. 7

The last result is suggested by Fig. 7. Here, the right triangle PQR is approximated by the figure PQS , PS being the arc of the circle with center at O and radius equal to r . The arc PS is therefore equal to $r\Delta\theta$, and the line SQ is equal to Δr . From the figure, it appears that

$$\tan \phi \text{ is approximated by } r \Delta\theta / \Delta r,$$

with the approximation becoming better as $\Delta\theta$ approaches 0. This is, of course, not a proof that $\lim_{\Delta\theta \rightarrow 0} \tan \phi = r d\theta/dr$. The latter result was proved above. The diagram in Fig. 7, leading intuitively to the fraction $r \Delta\theta / \Delta r$ as an approximate value of $\tan \phi$, thus suggests the value $r d\theta/dr$ for $\tan \psi$.

In the above discussion we thought of ψ as the angle OPT' , Fig. 6. Since this angle is equal to angle APT we can associate ψ with the latter angle. In so doing we shall agree to measure ψ counterclockwise from the radius vector produced beyond P , toward that part of the tangent PT which extends in the direction of increasing θ .

If OX extends horizontally toward the right from O , the angle ψ

between OX and the direction PT is the angle of inclination of the tangent, and $\tan \alpha$ is the slope. From Fig. 6, $\alpha = \theta + \psi$. Hence,

$$\begin{aligned}\text{slope of } PT &= \tan \alpha = \tan(\theta + \psi) \\ &= \frac{\tan \theta + \tan \psi}{1 - \tan \theta \tan \psi} \\ &= \frac{\tan \theta + r(d\theta/dr)}{1 - \tan \theta \cdot r(d\theta/dr)}.\end{aligned}$$

The same result is obtained if we use the general formula, $\alpha = \theta + \psi - n\pi$.

It is seldom necessary to make use of α , since any information about the direction of a polar curve that we may require is usually given by ψ .

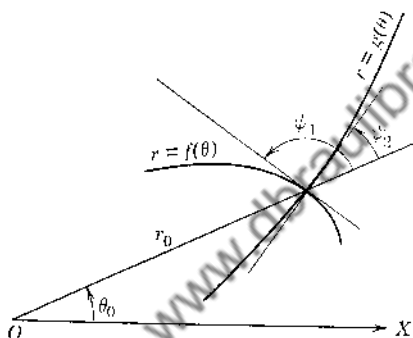


FIG. 8

For example, if two curves $r = f(\theta)$ and $r = g(\theta)$ intersect at a point P , their tangents meet at right angles if, and only if, $(\psi_1 - \psi_2) = \pm \pi/2$, where ψ_1 applies to the first curve and ψ_2 applies to the second. This is easily seen from Fig. 8. Hence

$$\tan \psi_1 = -\frac{1}{\tan \psi_2} \quad (3.1)$$

If the point P has the polar coordinates (r_0, θ_0) ,

$$\tan \psi_1 = \frac{r_0}{f'(\theta_0)} = \frac{f(\theta_0)}{f'(\theta_0)},$$

$$\tan \psi_2 = \frac{r_0}{g'(\theta_0)} = \frac{g(\theta_0)}{g'(\theta_0)}.$$

The equation $\tan \psi_1 = -\frac{1}{\tan \psi_2}$ then becomes

$$\frac{r_0}{f'(\theta_0)} = -\frac{1}{r_0/g'(\theta_0)}.$$

This relation is equivalent to the relation in rectangular coordinates

$$F'(x_0) = -1/G'(x_0),$$

when the graphs of $y = F(x)$ and $y = G(x)$ meet perpendicularly at (x_0, y_0) .

The formula for the angle of intersection of two polar curves is also similar to that in rectangular coordinates. As in Fig. 8 (where the angle of intersection is $\pi/2$), the angle of intersection is $\psi_1 - \psi_2$. Then,

$$\tan(\psi_1 - \psi_2) = \frac{\tan \psi_1 - \tan \psi_2}{1 + \tan \psi_1 \tan \psi_2}, \quad (3.2)$$

where

$$\tan \psi_1 = \frac{r_0}{f'(\theta_0)}, \quad \tan \psi_2 = \frac{r_0}{g'(\theta_0)}.$$

The student should compare formula 3.2 with formula 2.1, Chapter 14.

If the curve passes through the pole, at this point $\tan \psi$ is equal to 0. This follows from the definition of $\tan \psi$ as $\lim_{\Delta\theta \rightarrow 0} \tan \phi$. In the present case, θ is the value of the vectorial angle for which $r = 0$, that is, the value of θ for which $f(\theta) = 0$. Since $r = 0$, it follows that $r \sin \Delta\theta = 0$, and $r(1 - \cos \Delta\theta) + \Delta r = \Delta r$. Hence,

$$\tan \phi = \frac{r \sin \Delta\theta}{r(1 - \cos \Delta\theta) + \Delta r} = \frac{0}{\Delta r} = 0,$$

and therefore,

$$\lim_{\Delta\theta \rightarrow 0} \tan \phi = \tan \psi = 0.$$

Thus, the limiting direction of the tangent is the same as the direction of the radius vector as the point approaches the pole.

We consider some examples illustrating the above results.

EXAMPLE 1. Find $\tan \psi$ at any point of the curve $r = 10 \sin \theta$.

SOLUTION. Differentiating, $dr/d\theta = 10 \cos \theta$. Hence,

$$\tan \psi = \frac{r}{dr/d\theta} = \frac{10 \sin \theta}{10 \cos \theta} = \tan \theta.$$

This curve is the circle shown in Fig. 9.

EXAMPLE 2. Find $\tan \psi$ at any point of the curve $r = ae^\theta$. Find the slope at $r = ae$, $\theta = 1$.

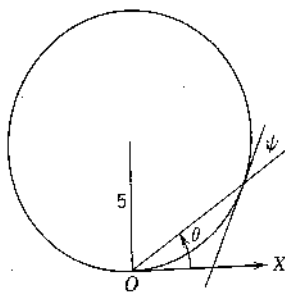


FIG. 9

SOLUTION. Differentiating, $dr/d\theta = ae^\theta$. Hence,

$$\tan \psi = \frac{r}{dr/d\theta} = \frac{ae^\theta}{ae^\theta} = 1,$$

and ψ is constantly = to $\pi/4$.

When $\theta = 1$, $\alpha = \psi + \theta = \frac{\pi}{4} + 1$.

Hence,

$$\tan \alpha = \tan (\pi/4 + 1) = -4.5880.$$

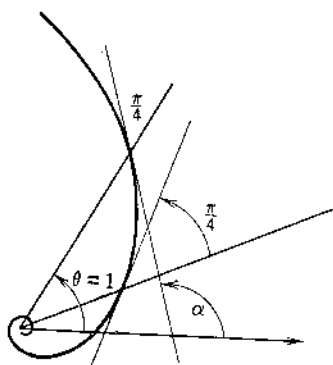


FIG. 10

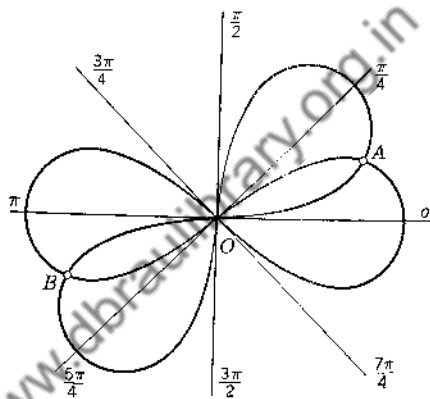


FIG. 11

As $\theta \rightarrow -\infty$, this curve, known as a logarithmic spiral, coils around the pole as indicated in Fig. 10.

EXAMPLE 3. Find the angles of intersection of the curves $r^2 = 4 \sin 2\theta$ and $r^2 = 4 \cos 2\theta$ (Fig. 11).

SOLUTION. One sees from the figure that there are three points of intersection: O , A , and B . Moreover, the angle of intersection at B is equal to that at A . Let us first find the points of intersection by solving the two equations simultaneously.

Setting $4 \sin 2\theta = 4 \cos 2\theta$, since both are equal to r^2 at the point of intersection, we have

$$\sin 2\theta = \cos 2\theta, \quad \text{or} \quad \tan 2\theta = 1,$$

whence,

$$2\theta = \pi/4, 5\pi/4, 9\pi/4, \text{ etc.} \quad \text{and} \quad \theta = \pi/8, 5\pi/8, 9\pi/8, \text{ etc.}$$

The point A is $\theta = \pi/8$, $r = \sqrt[4]{8}$. The value $\theta = 5\pi/8$, $2\theta = 5\pi/4$, does not correspond to a point of intersection since, for this value, in the second equation $r^2 < 0$ and r is imaginary. For $\theta = 9\pi/8$, $2\theta = 9\pi/4$, we have $r = \sqrt[4]{8}$. This is point B . Whereas the point O is a point of

intersection of the curves, it is not found by solving the equations simultaneously. The pole of the polar coordinate system has the peculiar property of corresponding to the coordinates $(0, \theta)$ where θ may be any value whatever. Thus, the point O lies on the curve $r^2 = 4 \sin 2\theta$, because this equation is satisfied by $r = 0, \theta = 0$, and also on the curve $r^2 = 4 \cos 2\theta$, which is satisfied by $r = 0, \theta = \pi/4$. Neither of these pairs of coordinates, however, is a solution of both equations.

At O there are eight angles of intersection of one curve with the other, consisting of four pairs of vertical angles. The four angles on the right-hand side of the vertical axis are shown in Fig. 12. The arcs MO and NO lie on $r^2 = 4 \sin 2\theta$, while PO and QO lie on $r^2 = 4 \cos 2\theta$. The angle between MO and PO is $\pi/4$ since MO enters the pole at $\theta = 0$, and PO enters it at $\theta = \pi/4$. Similarly, NO and PO intersect at an angle of $\pi/4$, and so do QO and MO , whereas NO and QO intersect at an angle of $3\pi/4$.

To find the angle of intersection at A we use (3.1) or (3.2), whichever is appropriate. Differentiating $r^2 = 4 \sin 2\theta$,

$$r \frac{dr}{d\theta} = 4 \cos 2\theta,$$

whence, by division,

$$\frac{r}{dr/d\theta} = \tan 2\theta.$$

Differentiating $r^2 = 4 \cos 2\theta$,

$$r \frac{dr}{d\theta} = -4 \sin 2\theta,$$

whence, by division,

$$\frac{r}{dr/d\theta} = -\cot 2\theta.$$

Hence, at $(\sqrt[4]{8}, \pi/8)$,

$$\tan \psi_1 = \tan 2 \left(\frac{\pi}{8} \right) = 1.$$

$$\tan \psi_2 = -\cot 2 \left(\frac{\pi}{8} \right) = -1.$$

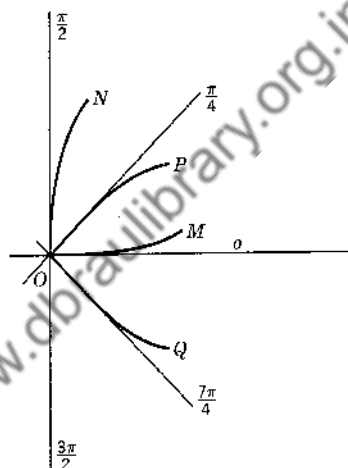


FIG. 12

Thus, $\tan \psi_1 = -\frac{1}{\tan \psi_2}$, and the curves meet perpendicularly. From the symmetry of the curves we see that the angle of intersection at B is also a right angle.

EXAMPLE 4. Find the angle of intersection of the cardioid $r = a(1 - \cos \theta)$ and the circle $r = 2a \cos \theta$ (Fig. 13).

SOLUTION. We first solve for the points of intersection by setting

$$2a \cos \theta = a(1 - \cos \theta).$$

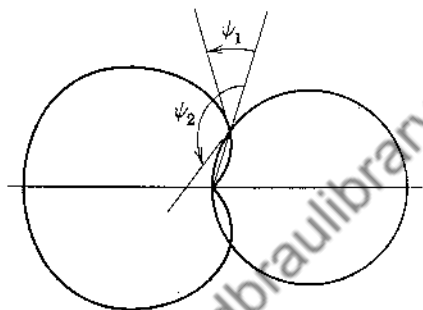


FIG. 13

From this equation we obtain

$$\cos \theta = \frac{1}{3}, \quad \theta = \cos^{-1} \frac{1}{3}, \quad r = \frac{2}{3}a.$$

Differentiating both equations, from the first we find

$$dr/d\theta = a \sin \theta,$$

$$\tan \psi_1 = \frac{r}{dr/d\theta} = \frac{a(1 - \cos \theta)}{a \sin \theta} = \frac{1 - 1/3}{(2\sqrt{2})/3} = \frac{1}{\sqrt{2}}.$$

From the second equation,

$$dr/d\theta = -2a \sin \theta,$$

$$\tan \psi_2 = \frac{r}{dr/d\theta} = \frac{2a \cos \theta}{-2a \sin \theta} = -\cot \theta = -\frac{1}{2\sqrt{2}}.$$

Putting these values in Eq. (3.2),

$$\tan (\psi_2 - \psi_1) = \frac{-1/(2\sqrt{2}) - 1/\sqrt{2}}{1 - 1/(2\sqrt{2}) \cdot (1/\sqrt{2})} = -\sqrt{2}.$$

Hence,

$$\psi_2 - \psi_1 = \tan^{-1}(-\sqrt{2}) = 125^\circ 16'.$$

EXERCISES

1. Find ψ and ϕ in terms of θ for the circle $r = a \cos \theta$.
2. Show that ψ is a constant for all points on the spiral $r = e^{a\theta}$, and find its value if $a = 3$. To what does this curve degenerate when $a = 0$, and what is the value of ψ in this case?
3. Find $\tan \psi$ for the cardioid $r = a(1 - \cos \theta)$ at $\theta = \pi/2$, $\theta = \pi$, and $\theta = 3\pi/2$. Find ψ at the pole.
4. The graph of $r = b - a \cos \theta$, with $a \neq b$, is called a *limaçon*. The equation $r = 2 - 4 \cos \theta$ is graphed in Fig. 14. Find the values of ψ at $(-2, 0)$, $(2, \pi/2)$, and $(6, \pi)$. At what angle does this limaçon intersect itself at the pole?

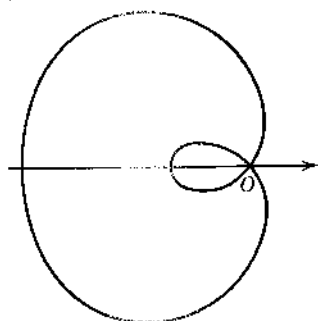


FIG. 14

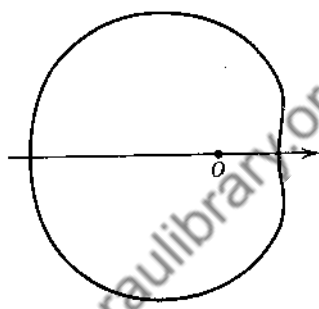


FIG. 15

5. Fig. 15 shows the limaçon $r = 4 - 2 \cos \theta$. Find ψ at $(2, 0)$, $(4, \pi/2)$, $(6, \pi)$, $(3, \pi/3)$.
6. At what angle does the cardioid $r = 4 - 4 \cos \theta$ intersect the limaçon $r = 4 - 2 \cos \theta$?
7. Show that for the parabola $r = a \sec^2 \theta/2$, $\alpha + \psi = \pi$.
8. Show that there is no value of b for which the circle $r = b$ intersects the lemniscate $r^2 = a^2 \sin 2\theta$ perpendicularly.
9. Find the slope of the spiral of Archimedes $r = a\theta$ at $(a, 1)$.
10. Find the slope of the reciprocal spiral $r = a/\theta$ at $(a, 1)$.
11. Show that wherever the two spirals of Exercises 9 and 10 intersect, their tangents are perpendicular.
12. Show that the following pairs of curves intersect perpendicularly. (a) The circles $r = 10 \sin \theta$ and $r = 10 \cos \theta$. (b) The cardioids $r = a(1 - \cos \theta)$ and $r = a(1 + \cos \theta)$. (c) The parabolas $r = a \sec^2 \theta/2$ and $r = a \csc^2 \theta/2$. (d) The curves $r^2 \theta = a^2$ and $r = e^{\theta^3}$.

4. **Arc Length in Polar Coordinates.** Formula (6.3) of Chapter 14, for the length of an arc, can be transformed so as to give a formula for arc length in polar coordinates. This formula is

$$s = \int_{t_1}^{t_2} \sqrt{(dx/dt)^2 + (dy/dt)^2} dt, \quad (4.1)$$

where the coordinates of points on the curve are expressed parametrically in terms of t . If the polar equation of some curve is

$$r = f(\theta),$$

we can express the rectangular coordinates, x and y , in terms of θ as a parameter as follows.

Since $x = r \cos \theta$, $y = r \sin \theta$, and on the curve $r = f(\theta)$, it follows that for all points on the curve

$$x = f(\theta) \cos \theta, \quad y = f(\theta) \sin \theta.$$

Differentiating these equations,

$$dx/d\theta = f'(\theta) \cos \theta - f(\theta) \sin \theta$$

$$dy/d\theta = f'(\theta) \sin \theta + f(\theta) \cos \theta.$$

Hence,

$$(dx/d\theta)^2 + (dy/d\theta)^2 = [f'(\theta)]^2 + [f(\theta)]^2.$$

Substituting this expression in (4.1),

$$\begin{aligned} s &= \int_{\theta_1}^{\theta_2} \sqrt{[f'(\theta)]^2 + [f(\theta)]^2} d\theta \\ &= \int_{\theta_1}^{\theta_2} \sqrt{(dr/d\theta)^2 + r^2} d\theta. \end{aligned} \quad (4.2)$$

If the length s is thought of as the variable length from some fixed point to a moving point, then s is a function of θ .

$$s = \int_{\theta_1}^{\theta} \sqrt{[f'(u)]^2 + [f(u)]^2} du$$

It follows from the fundamental theorem that

$$ds/d\theta = \sqrt{[f'(\theta)]^2 + [f(\theta)]^2} = \sqrt{(dr/d\theta)^2 + r^2} \quad (4.3)$$

and

$$ds = \sqrt{(dr/d\theta)^2 + r^2} d\theta. \quad (4.4)$$

EXAMPLE. Find the complete length of the cardioid

$$r = a(1 - \cos \theta).$$

SOLUTION. Differentiating, and substituting in (4.2), we have

$$\begin{aligned} s &= \int_0^{2\pi} \sqrt{a^2 \sin^2 \theta + a^2(1 - \cos \theta)^2} d\theta \\ &= \int_0^{2\pi} a \sqrt{2 \cdot 2 \sin^2 (\theta/2)} d\theta \\ &= 2a \int_0^{2\pi} \sin (\theta/2) d\theta = \left[-4a \cos (\theta/2) \right]_0^{2\pi} = 8a. \end{aligned}$$

EXERCISES

1. Find the length of the circumference of the circle $r = \sin \theta$ by integration in polar coordinates.
2. Find the length of the spiral $r = e^{a\theta}$ from $(1, 0)$ to $(e^{a\pi}, \pi)$.
3. Find the length of the spiral $r\theta = a$ from $(2a, \frac{1}{2})$ to $(a/\pi, \pi)$.
4. Find the length of the arc of the curve $r = 2a \tan \theta \sin \theta$ from $\theta = 0$ to $\theta = \pi/3$.
5. Obtain the formula

$$s = \int_{r_0}^{r_1} \sqrt{r^2(d\theta/dr)^2 + 1} \, dr \quad (4.5)$$

from (4.2).

6. Find the entire length of the curve $r = a \cos^3 \frac{\theta}{3}$ (Fig. 16).

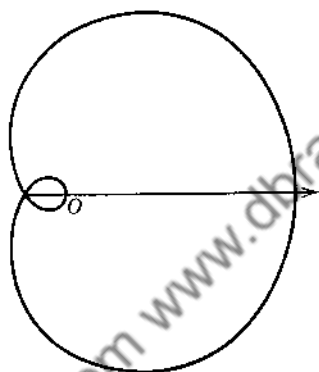


FIG. 16

5. Curvature in Polar Coordinates. The curvature of a curve was defined in Art. 8, Chapter 14, as the rate of change of direction with respect to distance along the curve. Thus, we write

$$K = d\alpha/ds,$$

where α is the angle which the tangent to the curve makes with a line of fixed direction. Just as in Chapter 14 we let the fixed direction be that of the positive x -axis, so here we can let the fixed direction be that of the polar axis. That is, $\alpha = \theta + \psi$. Denoting differentiation with respect to θ by primes, we write $d\alpha/ds = \alpha'/s'$. Since $s' = \sqrt{(r')^2 + r^2}$ by formula 4.3, we have

$$K = \frac{\alpha'}{\sqrt{(r')^2 + r^2}}. \quad (5.1)$$

Since $\alpha = \theta + \psi$,

$$\begin{aligned}\alpha' &= \frac{d}{d\theta}(\theta + \psi) = 1 + \frac{d}{d\theta}(\tan^{-1} r/r') \\&= 1 + \frac{1}{1 + r^2/(r')^2} \cdot \frac{d}{d\theta} \left(\frac{r}{r'} \right) \\&= 1 + \frac{1}{1 + r^2/(r')^2} \cdot \frac{(r')^2 - rr''}{(r')^2} \\&= 1 + \frac{(r')^2 - rr''}{(r')^2 + r^2} \\&= \frac{r^2 + 2(r')^2 - rr''}{(r')^2 + r^2}.\end{aligned}$$

Inserting this expression for α' in (5.1), we have

$$K = \frac{r^2 + 2(r')^2 - rr''}{[(r')^2 + r^2]^{\frac{3}{2}}}. \quad (5.2)$$

Since the radius of curvature, R , is equal to $1/|K|$,

$$R = \left| \frac{[(r')^2 + r^2]^{\frac{3}{2}}}{r^2 + 2(r')^2 - rr''} \right|.$$

EXAMPLE. Find the curvature of the lemniscate $r^2 = a^2 \cos 2\theta$ at any point. Draw the circle of curvature at the point $(a, 0)$.

SOLUTION. Differentiating $r^2 = a^2 \cos 2\theta$,

$$rr' = -a^2 \sin 2\theta, \quad r' = \frac{-a \sin 2\theta}{\sqrt{\cos 2\theta}},$$

$$rr'' + (r')^2 = -2a^2 \cos 2\theta.$$

Thus,

$$r^2 + 2(r')^2 - rr'' = 3a^2 \left(\cos 2\theta + \frac{\sin^2 2\theta}{\cos 2\theta} \right) = \frac{3a^2}{\cos 2\theta},$$

and

$$r^2 + (r')^2 = a^2 \left(\cos 2\theta + \frac{\sin^2 2\theta}{\cos 2\theta} \right) = \frac{a^2}{\cos 2\theta}.$$

Hence,

$$K = \frac{3\sqrt{\cos 2\theta}}{a}.$$

At the point $(a, 0)$, $K = 3/a$ and $R = a/3$. Figure 17 shows the circle of curvature at this point.

EXERCISES

1. Show that the curvature of the circle $r = 2a \cos \theta$ is constant.
2. Show that the radius of curvature of the curve $r = be^{a\theta}$ is proportional to the radius vector.
3. Find the curvature of the cardioid $r = a(1 - \cos \theta)$ at any point (r_1, θ_1) .

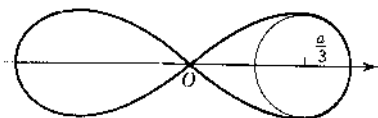


FIG. 17

4. Find the curvature of the spiral of Archimedes, $r = a\theta$, at any point (r_1, θ_1) .
5. Find the radius of curvature of the limaçon $r = 2 - 4 \cos \theta$ at the point $(6, \pi)$. Draw the curve and the circle of curvature at this point.

6. Area in Polar Coordinates. We studied the area problem that is suggested by figures on a rectangular coordinate system in Chapter 7. We shall now attack the area problem that is suggested by the representation of a function in polar coordinates. This approach will lead to a new definition of area, appropriate to figures on a polar coordinate system. The new definition is consistent with the earlier one, although this will not be proved. That is, if the area of a figure were evaluated by both definitions the results would agree.

The area problem suggested by a polar curve is the following. Let a curve in polar coordinates have the equation, $r = f(\theta)$. We require the area bounded by this curve and two radii vectors, OP and OQ (Fig. 18).

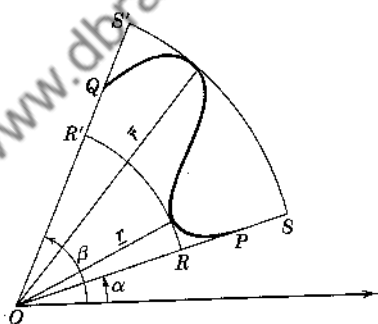


FIG. 18

Intuitively, we conceive this area as lying between two circular sectors, ORR' and OSS' in the illustration. The area of ORR' is $(\beta - \alpha)/2\pi$ times the area of the circle with radius r . Hence,

$$\text{area } ORR' = (\beta - \alpha)/2\pi \cdot \pi r^2 = \frac{1}{2}(\beta - \alpha)r^2.$$

Similarly,

$$\text{area } OSS' = \frac{1}{2}(\beta - \alpha)r'^2.$$

Moreover, as noted above, we are intuitively satisfied that

$$\frac{1}{2}(\beta - \alpha)r^2 \leq A \leq \frac{1}{2}(\beta - \alpha)r'^2. \quad (6.1)$$

If we now break up the given figure into n parts by drawing radial lines so that the angle $\beta - \alpha$ is subdivided into the parts $\Delta_1\theta$, $\Delta_2\theta$, etc., as

indicated in Fig. 19, we can place the area of each part between bounds in a similar fashion. Let $\Delta_i A$ denote the area of the part corresponding to $\Delta_i \theta$, and let $\underline{r}_i$ and $\overline{r}_i$ denote, respectively, the least and greatest value of r in this part. Then,

$$\frac{1}{2} \underline{r}_i^2 \Delta_i \theta \leq \Delta_i A \leq \frac{1}{2} \overline{r}_i^2 \Delta_i \theta.$$

By addition, and noting that $\sum_{i=1}^n \Delta_i A = A$, we have

$$\sum_{i=1}^n \frac{1}{2} \underline{r}_i^2 \Delta_i \theta \leq A \leq \sum_{i=1}^n \frac{1}{2} \overline{r}_i^2 \Delta_i \theta. \quad (6.2)$$

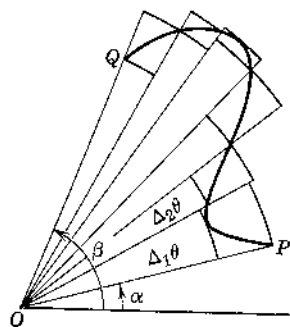


FIG. 19

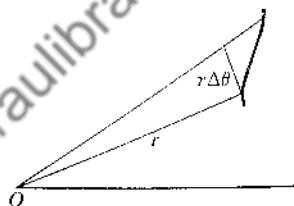


FIG. 20

The least of these sums is the sum of the areas of sectors lying wholly within the curve (on the side toward O), and the greatest is the sum of the areas of sectors extending beyond the curve.

The sums in (6.2) are exactly like those denoted by S_n and $\overline{S}_n$ in Art. 7, Chapter 7. Hence, if the manner of subdivision is such that each $\Delta_i \theta \rightarrow 0$ as $n \rightarrow \infty$, both sums approach the same limit. On account of (6.2), which shows that A always lies between the sums, it is natural to define the area A by this common limit. Thus,

$$A = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{2} r_i^2 \Delta_i \theta, \quad (6.3)$$

in which r_i is equal to $\underline{r}_i$ or $\overline{r}_i$ or, indeed, any value between these two.

This limit is the integral $\int_{\alpha}^{\beta} \frac{1}{2} r^2 d\theta$. Therefore,

$$A = \int_{\alpha}^{\beta} \frac{1}{2} r^2 d\theta = \frac{1}{2} \int_{\alpha}^{\beta} [f(\theta)]^2 d\theta. \quad (6.4)$$

(This definition is suggested by considering the area as the limit of the sum of triangles having altitude r and base $r \Delta \theta$, Fig. 20.)

EXAMPLE 1. Find the area bounded by the lemniscate in Fig. 21:

$$r^2 = a^2 \sin 2\theta.$$

SOLUTION. Since the lemniscate is made up of two equal parts, the whole area is

$$\begin{aligned} A &= 2 \int_0^{\pi/2} \frac{1}{2} a^2 \sin 2\theta \, d\theta \\ &= \left[-\frac{1}{2} a^2 \cos 2\theta \right]_0^{\pi/2} = a^2. \end{aligned}$$

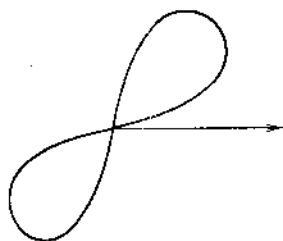


FIG. 21

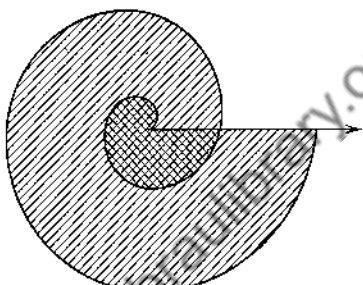


FIG. 22

EXAMPLE 2. Find the area swept out by the radius vector of the curve $r = e^\theta$, in two complete revolutions of the spiral about the pole (Fig. 22).

SOLUTION. We are tempted to write

$$A = \int_0^{4\pi} \frac{1}{2} e^{2\theta} \, d\theta = \left[\frac{1}{4} e^{2\theta} \right]_0^{4\pi} = \frac{1}{4} (e^{8\pi} - 1)$$

This is not correct, however, since the area swept out by the radius vector in its first revolution is included twice. Actually, the whole area is swept out in the second revolution. Hence

$$A = \int_{2\pi}^{4\pi} \frac{1}{2} e^{2\theta} \, d\theta = \left[\frac{1}{4} e^{2\theta} \right]_{2\pi}^{4\pi} = \frac{1}{4} (e^{8\pi} - e^{4\pi}) = \frac{1}{4} e^{4\pi} (e^{4\pi} - 1).$$

EXERCISES

1. Find the area enclosed by the lemniscate $r^2 = 16 \cos 2\theta$.
2. Find the area of all the loops of the curve $r = 4 \cos 2\theta$.
3. Find the area of all the loops of the curve $r = 2 \sin 3\theta$.
4. Find the area of the cardioid $r = a(1 - \cos \theta)$.
5. Find the area of the limaçon $r = 4 - 2 \cos \theta$.
6. Find the area of the inner loop of the limaçon $r = 2 - 4 \cos \theta$.
7. Find the area swept out by the radius vector of the curve $r = \frac{1}{2}\theta$ in two revolutions. Find the area enclosed between the first two coils of this spiral and bounded by the polar axis.

8. Find the area in the cardioid $r = 1 + \sin \theta$ and exterior to the circle $r = 1$.
9. Find the area in the circle $r = 4 \sin \theta$ and outside the circle $r = 2$.
10. Find the area common to the two circles $r = 2a \sin \theta$ and $r = 2a \cos \theta$.
11. Find the area in one loop of the lemniscate $r^2 = a^2 \cos 2\theta$ and outside the circle $r = a$.
12. Find the area included between the parabola $r = 1/(1 - \cos \theta)$ and the line $r \cos \theta = 8$ (Fig. 23). [Hint: $A = 2(OCA + OBA)$.]

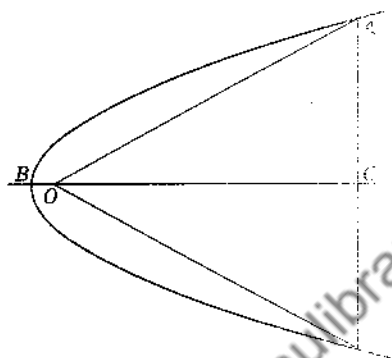


FIG. 23

13. Transform the equations in Exercise 12 into rectangular coordinates, and calculate the same area by use of the formula

$$A = \int_a^b y \, dx.$$

14. Find the area between the parabolas $r = 1/(1 - \cos \theta)$ and $r = 1/(1 + \cos \theta)$.
15. Find the area below the polar axis bounded by the curve $r = a \sin^3 \theta/3$ (see Fig. 16).

Physical Applications

1. Curvilinear Motion. Velocity in Polar Coordinates. Motion in a curved path was studied in Chapter 6. We shall carry that study further here. First, we shall summarize some of the earlier discussion. It will simplify notations throughout to denote derivatives with respect to t by dots. Thus,

$$dx/dt = \dot{x}, \quad dy/dt = \dot{y}, \quad ds/dt = \dot{s}, \quad d^2x/dt^2 = \ddot{x}, \quad dy/dx = \dot{y}/\dot{x}, \text{ etc.}$$

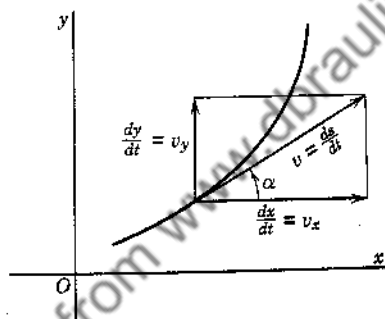


FIG. 1

If the coordinates of the position of a body are given by two functions of the time t ,

$$x = X(t), \quad y = Y(t),$$

the components of the velocity vector v are given by

$$v_x = \dot{x}, \quad v_y = \dot{y}, \quad \text{and} \quad |v| = \sqrt{v_x^2 + v_y^2}.$$

Also,

$$\dot{s} = \pm \sqrt{v_x^2 + v_y^2} = \pm \sqrt{\dot{x}^2 + \dot{y}^2},$$

in which the choice of the sign depends upon the choice of the positive direction along the curve. The vector velocity is determined by the components v_x and v_y . This vector is directed along the tangent, since $v_y/v_x = \dot{y}/\dot{x} = dy/dx = \tan \alpha$ (Fig. 1). Also, as illustrated in Fig. 1,

$$\dot{y}/\dot{s} = dy/ds = \sin \alpha, \quad \dot{x}/\dot{s} = dx/ds = \cos \alpha. \quad (1.1)$$

It is sometimes desirable to express the vector v by means of components, one of which lies in the direction of the radius vector and the other perpendicular to this direction (Fig. 2). This representation suggests the use of polar coordinates, and we denote these two components by v_r and v_θ , respectively. We consider the polar coordinates of points on the path to be given parametrically by

$$\theta = \Theta(t), \quad r = R(t).$$

Since v_r is the rate of motion in the direction of r , it is intuitively obvious that $v_r = \dot{r}$. Similarly, v_θ is the rate of motion in the direction of the

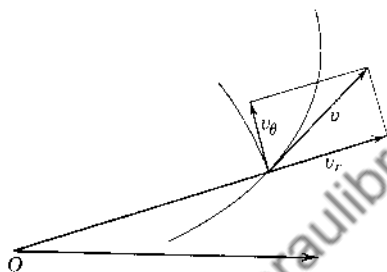


FIG. 2

circumference of a circle with radius r , and it is thus clear intuitively that v_θ is equal to the angular velocity $\dot{\theta}$ multiplied by the radius r . That is,

$$v_\theta = r\dot{\theta}. \quad (1.2)$$

From Fig. 2,

$$|v| = \sqrt{v_r^2 + v_\theta^2} = \sqrt{\dot{r}^2 + r^2\dot{\theta}^2}. \quad (1.3)$$

Hence,

$$\dot{s} = \pm \sqrt{\dot{r}^2 + r^2\dot{\theta}^2}. \quad (1.4)$$

From Chapter 15,

$$\tan \psi = r \frac{d\theta}{dr} = \frac{r\dot{\theta}}{\dot{r}} = \frac{v_\theta}{v_r}.$$

Hence,

$$\sin \psi = \frac{v_\theta}{\pm \sqrt{v_r^2 + v_\theta^2}} = \frac{r\dot{\theta}}{\pm \sqrt{\dot{r}^2 + r^2\dot{\theta}^2}} \quad (1.5)$$

$$= \frac{r\dot{\theta}}{\dot{s}} = r \frac{d\theta}{ds}. \quad (1.6)$$

It also follows that

$$v_\theta = \dot{s} \sin \psi. \quad (1.7)$$

Similarly,

$$\cos \psi = \frac{v_r}{\pm \sqrt{v_r^2 + v_\theta^2}} = \frac{\dot{r}}{\pm \sqrt{\dot{r}^2 + r^2 \dot{\theta}^2}} \quad (1.8)$$

$$= \frac{\dot{r}}{\dot{s}} = \frac{dr}{ds} \quad (1.9)$$

and

$$v_r = \dot{s} \cos \psi. \quad (1.10)$$

In these formulas, the sign of the radical must be chosen consistently with that of $d\theta/ds$. It is assumed that r is positive, and therefore $\sin \psi$

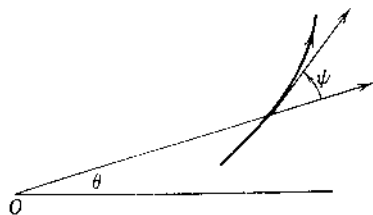


FIG. 3

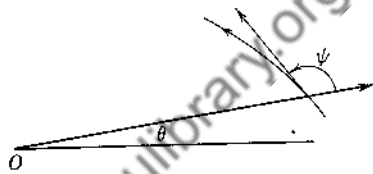


FIG. 4

must agree in sign with $d\theta/ds$. Hence, if s increases with θ , ψ must be an angle between 0 and π (Figs. 3 and 4). If s decreases as θ increases, ψ must be an angle of the third or fourth quadrants (Figs. 5 and 6).

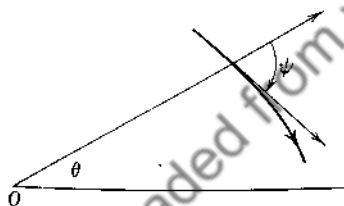


FIG. 5

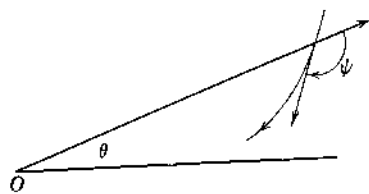


FIG. 6

EXAMPLE 1. A point moves so that its polar coordinates at any time t are given by the equations

$$r = 2(1 + \sin 5t), \quad \theta = 5t.$$

Find v_r , v_θ , and the magnitude of v at any time t . Find these values when $t = \pi/15$.

SOLUTION. Differentiating, we have

$$\dot{r} = v_r = 10 \cos 5t, \quad \dot{\theta} = 5.$$

Since

$$v_\theta = r\dot{\theta}, \text{ we have } v_\theta = 10(1 + \sin 5t).$$

From (1.3),

$$\begin{aligned} |v| &= \sqrt{100 \cos^2 5t + 100(1 + 2 \sin 5t + \sin^2 5t)} \\ &= \sqrt{200 + 200 \sin 5t} = 10\sqrt{2} \sqrt{1 + \sin 5t}. \end{aligned}$$

Putting $t = \pi/15$, $5t = \pi/3$, in these results, we obtain

$$v_r = 10 \cos (\pi/3) = 5,$$

$$v_\theta = 10[1 + \sin (\pi/3)] = 10(1 + \sqrt{3}/2),$$

$$|v| = 10\sqrt{2} \sqrt{1 + \sqrt{3}/2} = 10\sqrt{2 + \sqrt{3}}.$$

EXAMPLE 2. A point moves along the spiral $r = \frac{1}{2}\theta$ at a constant speed $\dot{s} = 2$ units per sec. Find v_r and v_θ at any point (Fig. 7).

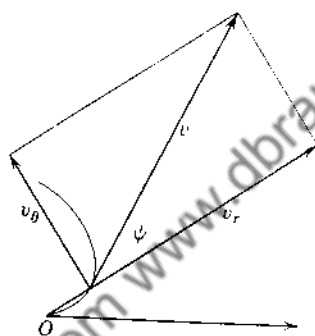


FIG. 7

SOLUTION. From the given equation, $dr/d\theta = \frac{1}{2}$, and

$$\tan \psi = \frac{r}{dr/d\theta} = \frac{\theta/2}{1/2} = \theta.$$

From Eqs. 1.7 and 1.10,

$$v_\theta = 2 \sin \psi, \quad \text{and} \quad v_r = 2 \cos \psi.$$

Since $\tan \psi = \theta$,

$$\sin \psi = \frac{\theta}{\sqrt{1 + \theta^2}}, \quad \cos \psi = \frac{1}{\sqrt{1 + \theta^2}}.$$

Hence,

$$v_\theta = \frac{2\theta}{\sqrt{1 + \theta^2}}, \quad v_r = \frac{2}{\sqrt{1 + \theta^2}}.$$

EXAMPLE 3. A point moves counterclockwise on the cardioid $r = 3(1 - \cos \theta)$, at a constant speed of 3 units per sec. Find v_r , v_θ , and the angular velocity at $\theta = \pi/2$, and $\theta = \pi$ (Fig. 8).

SOLUTION. We have $dr/d\theta = 3 \sin \theta$. Hence,

$$\tan \psi = \frac{r}{dr/d\theta} = \frac{3(1 - \cos \theta)}{3 \sin \theta} = \frac{1 - \cos \theta}{\sin \theta} = \tan \frac{\theta}{2}.$$

$$\sin \psi = \sin \frac{\theta}{2}, \quad \cos \psi = \cos \frac{\theta}{2}.$$

From Eqs. 1.7 and 1.10,

$$v_\theta = 3 \sin \frac{\theta}{2}, \quad v_r = 3 \cos \frac{\theta}{2}.$$

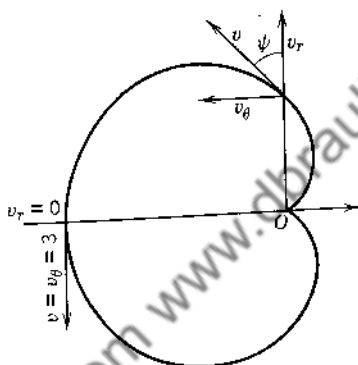


FIG. 8

The angular velocity $\dot{\theta}$ is

$$\begin{aligned} \dot{\theta} &= \frac{v_\theta}{r} = \frac{3 \sin (\theta/2)}{3(1 - \cos \theta)} \\ &= \frac{\sqrt{(1 - \cos \theta)/2}}{1 - \cos \theta} = \frac{1/2}{\sqrt{(1 - \cos \theta)/2}} \\ &= \frac{1}{2 \sin (\theta/2)} = \frac{\csc (\theta/2)}{2}. \end{aligned}$$

At $\theta = \pi/2$,

$$\begin{aligned} v_\theta &= 3 \sin (\pi/4) = 3\sqrt{2}/2, \quad v_r = 3 \cos (\pi/4) = 3\sqrt{2}/2, \\ \dot{\theta} &= \frac{1}{2} \csc (\pi/4) = \frac{1}{2}\sqrt{2}. \end{aligned}$$

At $\theta = \pi$,

$$v_\theta = 3 \sin (\pi/2) = 3, \quad v_r = 3 \cos (\pi/2) = 0, \quad \dot{\theta} = \frac{1}{2} \csc (\pi/2) = \frac{1}{2}.$$

It is interesting to observe that at $\theta = 0$ (or $\theta = 2\pi$), $\dot{\theta} = (\csc 0)/2$, which does not exist. This point, however, is a cusp of the cardioid, and the moving point is stationary, which is a contradiction to the requirement that it shall move at a constant speed of 3 units per sec. It must therefore be expected that the results obtained above need not hold at this point. In order to gain a clearer idea of the behavior of the moving point as it passes through the cusp, we may observe that the cardioid is approached more and more closely by the limaçon

$$r = 3(1 + \varepsilon - \cos \theta) \quad (1.11)$$

as ε approaches 0 (Fig. 9). As the limaçon passes into the cardioid there appear to be two factors governing the angular velocity: since the radius vector becomes small it must turn rapidly as the point passes through the vertical arc at $\theta = 0$, but the arc with a vertical tangent ceases to exist as soon as the limaçon becomes the cardioid with $\varepsilon = 0$. At the cusp of the cardioid there is no angular velocity and the result $\dot{\theta} = (\csc 0)/2$ is meaningless. The angular velocity of the point moving with a constant speed along the limaçon (1.11) at θ is found to be $1/\varepsilon$, (see Exercise 3), and, hence, as ε approaches 0, this angular velocity increases without limit.

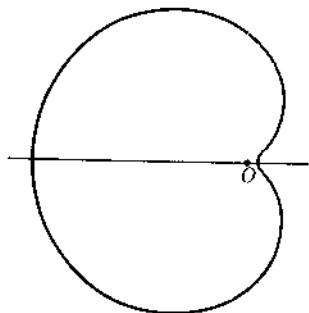


FIG. 9

EXERCISES

1. Find v_r , v_θ , and $|v|$ for the given values of t .

(a) $r = \sin t$, $\theta = 2t$; $t = \pi/3$, $t = \pi/2$.

(b) $r = 10t$, $\theta = t/2$; $t = 0$, $t = 3$.

(c) $r = 4 - 2 \cos 2t$, $\theta = 2t$; $t = 0$, $t = \pi/4$, $t = \pi/2$.

2. Find v_r , v_θ , and the angular velocity of a point moving counterclockwise with a constant speed = 10 along each of the following curves at the points indicated.

(a) $r = 10 \cos \theta$; $\theta = 0$, $\theta = \pi/6$, $\theta = \pi/2$.

(b) $r = ae^{b\theta}$; $\theta = 0$, $\theta = \theta_1$.

(c) $r\theta = 10$; $\theta = 1$, $\theta = \pi$.

(d) $r = 1/(1 - \cos \theta)$; $\theta = \pi/2$, $\theta = \pi$.

(e) $r = 4 - 2 \cos \theta$; $\theta = 0$, $\theta = \pi/2$, $\theta = \pi$.

3. Show that the angular velocity of the point moving with constant velocity on the limaçon $r = 3(1 + \varepsilon - \cos \theta)$ is $1/\varepsilon$ when $\theta = 0$.

2. Curvilinear Motion. Acceleration. In Chapter 6 we found that if the equations of motion are $x = X(t)$, $y = Y(t)$, the acceleration vector a is determined by the components

$$a_x = \ddot{x}, \quad a_y = \ddot{y}. \quad (2.1)$$

The magnitude of a is

$$|a| = \sqrt{a_x^2 + a_y^2} = \sqrt{\ddot{x}^2 + \ddot{y}^2}, \quad (2.2)$$

and the slope of a is a_y/a_x . The direction of a is not, in general, tangent to the path.

In a rectilinear motion the acceleration is the rate of change of velocity, which can only be a speeding up or a slowing down. In curvilinear motion, velocity can change by speeding up or slowing down in the path, or by changing direction, or by both. In complete analogy with the rectilinear case, the acceleration vector measures the change of the velocity, both in magnitude and direction. In order to study the effect of the acceleration vector, it is desirable to have it resolved into two components, one of which is tangential to the path and affects only the speed of the body, and the other of which is normal to the path and has no effect on the speed but only changes the direction of the motion. The first of these, denoted by a_t , will be found to be equal to ds/dt or $\dot{s}$. The second, denoted by a_n , is closely related to the curvature of the path. It is not equal to the curvature, since the rate of change of direction of the path depends not only on the normal component of acceleration, a_n , but also on the speed at which the body moves. We shall find that the curvature K is equal to $a_n/|v|^2$, or $a_n = |v|^2 K$.

We proceed to resolve a , whose components a_x and a_y we suppose to be known, into the components a_t and a_n (Fig. 10). For a_x we have

$$(\text{tangential component of } a_x) = a_x \cos \alpha,$$

$$(\text{normal component of } a_x) = a_x \cos (\alpha + 90^\circ) \\ = -a_x \sin \alpha.$$

Similarly, for a_y ,

$$(\text{tangential component of } a_y) = a_y \cos (90^\circ - \alpha) = a_y \sin \alpha,$$

$$(\text{normal component of } a_y) = a_y \cos \alpha.$$

Since a_t is equal to the sum of the two tangential components above, we have

$$a_t = a_x \cos \alpha + a_y \sin \alpha. \quad (2.3)$$

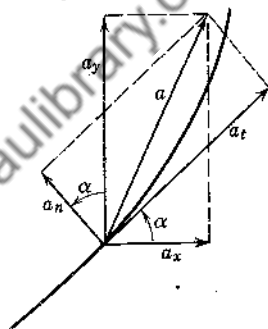


FIG. 10

In a similar way, we find

$$a_n = -a_x \sin \alpha + a_y \cos \alpha. \quad (2.4)$$

Equations 2.3 and 2.4 were derived in the same way that the equations for rotation of rectangular coordinate axes are found in analytic geometry. In particular, the positive directions on the tangent and normal are oriented in the same way as the positive x -direction and the positive y -direction. That is, the positive direction on the normal is 90° from that of the tangent, measured counterclockwise. The positive direction on the tangent depends on the choice of positive description of the curve.

Inserting $a_x = \ddot{x}$, $a_y = \ddot{y}$, $\cos \alpha = dx/ds$, $\sin \alpha = dy/ds$, in Eqs. 2.3 and 2.4, we obtain

$$a_t = \ddot{x} \frac{dx}{ds} + \ddot{y} \frac{dy}{ds} \quad (2.5)$$

$$a_n = -\ddot{x} \frac{dy}{ds} + \ddot{y} \frac{dx}{ds}. \quad (2.6)$$

From physical considerations, we are led to think that

$$a_t = ds/dt = \dot{s}. \quad (2.7)$$

Calculating $\dot{s}$ from $\dot{s} = \sqrt{\dot{x}^2 + \dot{y}^2}$ we obtain

$$\begin{aligned} \dot{s} &= \frac{\dot{x}\dot{x} + \dot{y}\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = \frac{\dot{x}\dot{x} + \dot{y}\dot{y}}{\dot{s}} \\ &= \dot{x} \frac{\dot{x}}{\dot{s}} + \dot{y} \frac{\dot{y}}{\dot{s}} = \dot{x} \frac{dx}{ds} + \dot{y} \frac{dy}{ds}. \end{aligned}$$

Since, by (2.5), this is equal to a_t , Eq. 2.7 is verified.

In order to compare the value of a_n with that of K , we write from Eq. 8.3, Chapter 14,

$$\begin{aligned} K &= \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{(\sqrt{\dot{x}^2 + \dot{y}^2})^3} = \frac{\dot{x}\ddot{y} - \ddot{x}\dot{y}}{\dot{s}^3} \\ &= \frac{1}{\dot{s}^2} \left(\ddot{y} \frac{\dot{x}}{\dot{s}} - \ddot{x} \frac{\dot{y}}{\dot{s}} \right) = \frac{1}{|v|^2} \left(\ddot{y} \frac{dx}{ds} - \ddot{x} \frac{dy}{ds} \right). \end{aligned}$$

Thus, from (2.6),

$$K = \frac{1}{|v|^2} a_n,$$

and

$$a_n = |v|^2 K. \quad (2.8)$$

Expressed in terms of the radius of curvature, R ,

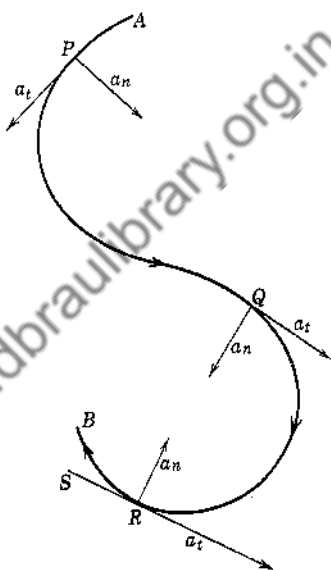
$$|a_n| = |v|^2/R. \quad (2.9)$$

In this equation, we use the absolute value of a_n , since $R = 1/|K|$ and the function $|v|^2/R$ is always positive. In Eq. 2.8, K , and consequently a_n , can be either positive or negative.

From physical considerations, it appears that a_n must be directed toward the concave side of the curve since that is the direction toward which the velocity vector is turned. This is also seen by means of Eq. 2.8. In Chapter 14 it was pointed out that K is positive at a point where the curve bends in a counterclockwise direction as one proceeds in the positive direction along the curve, and negative where the curve turns in a clockwise direction as one proceeds positively along the curve. By (2.8), the same statement would hold for a_n . Since the positive direction on the normal along which a_n is measured is always oriented 90° counterclockwise from the positive direction of the curve and the tangent to the curve, it follows that a_n always points toward the concave side of the curve. These conclusions are illustrated in Fig. 11, where we think of a particle moving along the curve from A toward B , the arrow heads indicating the positive description of the curve. At P , both a_t and a_n are positive, so that the particle is speeding up in its path, and the path is curving in the counterclockwise direction. The curvature is positive at this point. At Q , a_t is positive and a_n is negative, so that the particle is speeding up and the path curves in the clockwise direction. Note, also, that K has changed sign in passing from P to Q , and is now negative. At R , a_t is negative, and hence the particle is slowing in its path. The path continues to curve in the same way as at Q , i.e., clockwise, and a_n and K are still negative.

FIG. 11

In physics we learn that an acceleration takes place in the direction of a force and that its amount is proportional to the magnitude of the force. Thus, a_n is proportional to a force directed along the normal to the path toward the center of curvature, often called the *centripetal* force. This force must be transmitted to the moving body by some means. In a rotating wheel, by the tension in the spokes; for a truck going around a



bend, by the friction of the tires on the road; for a planet revolving around the sun, by whatever means gravitational forces are transmitted. It is worthy of note that, because of Eq. 2.8, this force varies directly as the square of the speed. For example, if the speed of a wheel is doubled, a_n is multiplied by 4 and so is the centripetal force. This rapid increase of centripetal force must be taken into account when the speed of rotation is greatly increased. Grindstones have been known to fly apart if run at too great a speed.

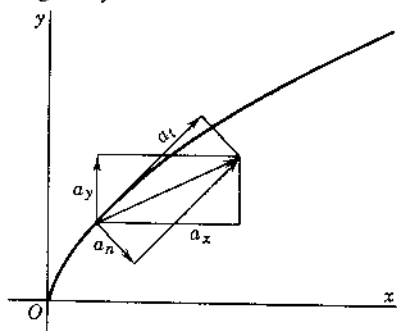


FIG. 12

EXAMPLE 1. A particle moves so that $x = t^3/3$, $y = t^3/2$, where t represents the time in seconds and x and y are measured in feet. Find expressions for a_t and a_n . Find a_t and a_n when $t = 1$.

SOLUTION. These equations of motion were given in Example 3, Art. 8, Chapter 6, where we found v_x , v_y , a_x , a_y , $|v|$, and $|a|$. We use Eqs. 2.5 and 2.6 to find a_t and a_n .

We have $\dot{x} = t^2$, $\dot{y} = t$, $\dot{s} = t\sqrt{t^2 + 1}$. If we assume that s increases with t we must choose the positive sign of the radical. Now,

$$dx/ds = \dot{x}/\dot{s} = t/\sqrt{t^2 + 1}, \quad dy/ds = \dot{y}/\dot{s} = 1/\sqrt{t^2 + 1},$$

$$\ddot{x} = 2t,$$

$$\ddot{y} = 1.$$

Hence, from (2.5) and (2.6),

$$a_t = 2t \cdot \frac{t}{\sqrt{t^2 + 1}} + 1 \cdot \frac{1}{\sqrt{t^2 + 1}} = \frac{2t^2 + 1}{\sqrt{t^2 + 1}},$$

and

$$a_n = -2t \cdot \frac{1}{\sqrt{t^2 + 1}} + 1 \cdot \frac{t}{\sqrt{t^2 + 1}} = -\frac{t}{\sqrt{t^2 + 1}}.$$

When

$$t = 1, \quad \sqrt{t^2 + 1} = \sqrt{2}, \quad \text{and} \quad a_t = 3/\sqrt{2}, \quad a_n = -1/\sqrt{2}.$$

The components of a are illustrated in Fig. 12.

EXAMPLE 2. A particle moves counterclockwise around a circle of radius r at a uniform speed u . Discuss the velocity and acceleration vectors.

SOLUTION. Placing the origin of the rectangular coordinate system at the center of the circle, the equations of motion are

$$x = r \cos \omega t, \quad y = r \sin \omega t, \quad (2.10)$$

in which ω is the constant angular velocity. The relation between ω and u is found from the velocity obtained from Eqs. 2.10. Thus,

$$v_x = -r\omega \sin \omega t, \quad v_y = r\omega \cos \omega t,$$

$$|v| = u = \sqrt{r^2\omega^2 \sin^2 \omega t + r^2\omega^2 \cos^2 \omega t} = r\omega,$$

and $\omega = u/r$. Hence, Eqs. 2.10 can be written

$$x = r \cos \frac{u}{r} t, \quad y = r \sin \frac{u}{r} t.$$

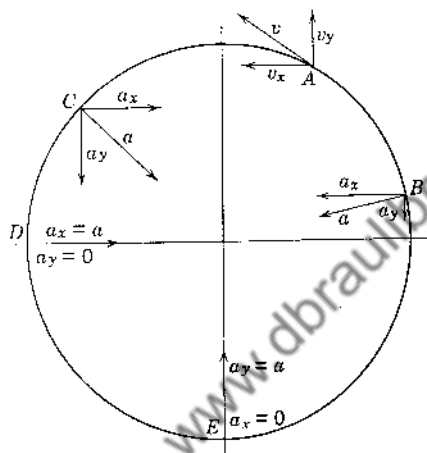


FIG. 13

The components of the acceleration in the direction of the axes are

$$a_x = \ddot{x} = -r\omega^2 \cos \omega t = -\omega^2 x,$$

$$a_y = \ddot{y} = -r\omega^2 \sin \omega t = -\omega^2 y.$$

Hence, a_x and a_y are proportional to x and y , respectively, and have signs opposite to these quantities. The tangential component of acceleration is 0, since $|v| = u = \text{constant}$. From Eq. 2.8,

$$a_n = u^2 K = u^2/r.$$

Here we consider the positive direction on the curve as that around the circle in the counterclockwise direction, agreeing with the direction of motion of the particle. Since $a_t = 0$, the acceleration vector a is directed toward the center and its magnitude is constantly u^2/r . Figure 13 shows several positions of the moving particle. At A the components of velocity are shown, and the acceleration is illustrated at B, C, D, and E.

EXAMPLE 3. A particle moves on an arc of the parabola $y^2 = 8x$, in the first quadrant, so that v_x is always equal to $10 - x$. Find expressions for v_y , $|v|$, a_t , and a_n at any point, and calculate these quantities at the instant when the particle passes through $(2, 4)$.

SOLUTION. Differentiating $y^2 = 8x$ with respect to t , we have $2y\dot{y} = 8\dot{x}$. That is, $2yv_y = 8v_x$, or $yv_y = 4v_x$. Since $v_x = 10 - x$, we have

$$yv_y = 4(10 - x), \quad v_y = 4(10 - x)/y.$$

Hence,

$$\begin{aligned} |v| &= \sqrt{v_x^2 + v_y^2} = (10 - x)\sqrt{1 + 16/y^2} \\ &= (10 - x)\sqrt{1 + 2/x}. \end{aligned}$$

Assuming that the particle moves so that s increases with t ,

$$\dot{s} = (10 - x)\sqrt{1 + 2/x}.$$

Differentiating this expression with respect to t ,

$$\begin{aligned} \dot{s} &= a_t = \left[(10 - x) \frac{-1/x^2}{\sqrt{1 + 2/x}} - \sqrt{1 + 2/x} \right] \dot{x} \\ &= \frac{(10 - x)(-x^2 - x - 10)}{x\sqrt{x}\sqrt{2 + x}}. \end{aligned}$$

In order to find a_n we use the formula $a_n = |v|^2 K$, and find K from formula 8.2, Chapter 14. From the given equation

$$dx/dy = \frac{1}{4}y, \quad d^2x/dy^2 = \frac{1}{4}.$$

Hence,

$$\begin{aligned} K &= \frac{d^2x/dy^2}{-[1 + (dx/dy)^2]^{3/2}} = -\frac{1/4}{(1 + y^2/16)^{3/2}} \\ &= -\frac{16}{(16 + 8x)^{3/2}} = -\frac{1}{\sqrt{2}(2 + x)^{3/2}}. \end{aligned}$$

The negative sign is chosen in order to make K negative in accordance with the convention making negative curvature correspond to clockwise turning. Therefore, employing the value of $|v|$ obtained above, we have

$$\begin{aligned} a_n &= -(10 - x)^2 \left(1 + \frac{2}{x}\right) \frac{1}{\sqrt{2}(2 + x)^{3/2}} \\ &= -\frac{(10 - x)^2}{x\sqrt{2}\sqrt{2 + x}}. \end{aligned}$$

At (2, 4),

$$v_x = 10 - 2 = 8, \quad v_y = 4(10 - 2)/4 = 8,$$

$$|v| = 8\sqrt{1+1} = 8\sqrt{2},$$

$$a_t = \frac{(10-2)(-4-2-10)}{2\sqrt{2}\sqrt{2+2}} = -16\sqrt{2}$$

$$a_n = -\frac{(10-2)^2}{2\sqrt{2}\sqrt{2+2}} = -8\sqrt{2}.$$

Alternatively, we can proceed from v_x and v_y to a_x and a_y , and then find a_t and a_n by means of formulas 2.5 and 2.6.

EXERCISES

1. For each of the pairs of parametric equations of motion given, find expressions for a_x , a_y , a_t , a_n , and evaluate these quantities for the indicated values of t .

(a) $x = 2t$, $y = t^2$; $t = 2$.

(b) $x = 300t$, $y = 400t - 16t^2$; $t = 12.5$, $t = 25$.

(c) $x = 10 \cos t$, $y = 6 \sin t$; $t = \pi/3$, $t = \pi/2$.

(d) $x = 5 \cos 2t$, $y = 5 \sin 2t$; $t = 0$, $t = \pi/4$, $t = \pi/2$.

(e) $x = e^t$, $y = e^{-t}$; $t = 0$, $t = 1$.

2. A particle moves around the curve $x = 4 \cos t + \cos 4t$, $y = 4 \sin t - \sin 4t$. Find the speed and the tangential acceleration at any time t .

3. A wheel with a 2-ft radius is rolling along a flat surface at 30 mph. Discuss the motion of a point on its rim. In particular, find the tangential and normal acceleration at any time, and the point at which the normal acceleration is maximum. [Hint: The curve described by the moving point is a cycloid with parametric equations,

$$x = a(\theta - \sin \theta), \quad y = a(1 - \cos \theta).$$

See Example 3, Art. 5, Chapter 14. Here $\theta = 22t$ where t denotes the time in seconds.]

4. A particle moves on an arc of the parabola $12y = x^2$ in the first quadrant, in such a way that v_y is always 8. Find expressions for v_x , $|v|$, a_x , a_y , a_t , and a_n . Draw a diagram showing a_t , a_n , and the vector a at the point (6, 3).

5. A particle moves on the hyperbola $x^2 - y^2 = 36$. If x and y are measured in feet, and v_x is always 8 ft per sec, find v_x , $|v|$, a_t , and a_n when the particle moves through (10, 8).

3. Acceleration. Polar Coordinates. We saw how to express the velocity in polar coordinates by means of formula 1.4. The tangential acceleration, a_t , can now be found from

$$a_t = \frac{d}{dt} \left(\frac{ds}{dt} \right) = \dot{s}.$$

The normal component of acceleration can be found as in Example 2, Art. 2, from the formula

$$a_n = |v|^2 K. \quad (2.8)$$

EXAMPLE 1. A particle moves on the spiral $r = \frac{1}{2}\theta$ with constant angular velocity $\dot{\theta} = 2$ rad per min. Find a_t and a_n when $\theta = 2$ rad.

SOLUTION. Differentiating $r = \frac{1}{2}\theta$,

$$\dot{r} = \frac{1}{2}\dot{\theta} = 1.$$

Hence,

$$s = \sqrt{r^2 + r^2\dot{\theta}^2} = \sqrt{1 + \frac{1}{4}\theta^2 \cdot 4} = \sqrt{1 + \theta^2},$$

$$a_t = \dot{s} = \frac{\theta \dot{\theta}}{\sqrt{1 + \theta^2}} = \frac{2\theta}{\sqrt{1 + \theta^2}}.$$

In order to find a_n , we calculate K from formula 5.2, Chapter 15.

$$K = \frac{r^2 + 2(r')^2 - r r''}{[r^2 + (r')^2]^{\frac{3}{2}}}.$$

In this formula, putting $r = \frac{1}{2}\theta$, $r' = \frac{1}{2}$, $r'' = 0$, we have

$$K = \frac{\frac{1}{4}\theta^2 + \frac{1}{2}}{(\frac{1}{4}\theta^2 + \frac{1}{4})^{\frac{3}{2}}} = \frac{2\theta^2 + 4}{(\theta^2 + 1)^{\frac{3}{2}}}.$$

Now, applying formula 2.8,

$$a_n = (1 + \theta^2) \frac{2\theta^2 + 4}{(1 + \theta^2)^{\frac{3}{2}}} = \frac{2\theta^2 + 4}{\sqrt{1 + \theta^2}}.$$

When $\theta = 2$,

$$a_t = \frac{2 \cdot 2}{\sqrt{1 + 2^2}} = \frac{4}{\sqrt{5}},$$

$$a_n = \frac{2 \cdot 4 + 4}{\sqrt{2^2 + 1}} = \frac{12}{\sqrt{5}}.$$

EXERCISES

1. A particle moves on the circle $r = 10 \sin \theta$ at a constant speed of 10 units. Find a_t and a_n when $\theta = \pi/4$.
2. A particle moves on the spiral $r = e^\theta$ with constant radial velocity, $v_r = 1$. Find $|v|$, and a_t and the angular velocity when $\theta = 0$.
3. A particle moves on the cardioid $r = 1 - \cos \theta$ with constant angular velocity, $\dot{\theta} = 2$. Find $|v|$, a_t , and a_n when $\theta = \pi/2$.

4. A New Work Problem. The work problem, first described in Chapter 1 as the physical counterpart of the area problem, was later discussed in some detail in Chapter 9. It will be remembered that work is found by multiplying a force by a distance. The feature of this work problem that took it out of the range of elementary algebraic methods, and required integration for its solution, was the variability of the force. In a new work problem that we shall now take up, it is the variability of the distance

that requires the application of the operation of integration. We shall consider an example.

EXAMPLE. A cistern 8 ft square and 6 ft deep is full of water. Find the amount of work done in pumping the water out, i.e., in bringing all the water to the top.

SOLUTION. We shall assume that water weighs 62.5 lb per cu ft. This means that a force of 62.5 lb must be applied to each cubic foot of water to lift it. If a certain cubic foot is raised 2 ft, $2 \cdot 62.5 = 125$ ft-lb, of work is done. But some of the water in the cistern has to be raised 6 ft and some not at all. All of it must be raised some distance between 0 and 6 ft. Of course, if a pipe were inserted all the way to the bottom and a pump attached, all the water would be brought through this pipe. One might then reason that all of it is raised 6 ft and conclude that the amount of work is equal to the product of the weight of all the water multiplied by 6. A little reflection

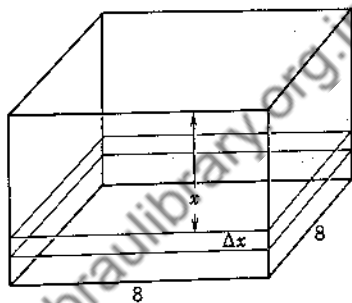


FIG. 14

shows that this is not correct, since each cubic inch of water that goes down to the bottom in order to enter the pipe pushes another cubic inch upward through the pipe to the height from which it started. The total amount of work is the same as though the end of the pipe were continually lowered just fast enough to keep it at the surface.

We solve this problem in much the same way that we solved the problem of total pressure on a vertical wall, in Chapter 9. We consider a layer of water of thickness Δx and lying between the depths x and $x + \Delta x$ (Fig. 14). Its volume is $8 \cdot 8 \cdot \Delta x$ cu ft, and its weight is $8 \cdot 8 \cdot \Delta x \cdot 62.5$ lb. The force required to lift this layer is therefore $8 \cdot 8 \cdot 62.5 \Delta x = 4000 \Delta x$, and the work done in raising it to the top is approximately $4000 \Delta x \cdot x$. The total amount of work done in raising all the water to the top is given by the integral,

$$W = \int_0^6 4000 x \, dx = 72,000.$$

The total work done is therefore 72,000 ft-lb.

EXERCISES

1. If the cistern in the illustrative example is one-half full, find the work done in emptying it over the top.
2. Find the work done in pumping the water in a cylindrical cistern to the top, if the cistern is 6 ft in radius and 12 ft deep and is full.

3. A conical tank is 12 ft across the top and 10 ft deep. If it is full of water find the work done in pumping it to the top.
4. If the water in the tank in Exercise 3 is 8 ft deep at its deepest point, find the work done in pumping the water to the top.
5. If the tank in Exercise 3 is full, find the work done in pumping the water to a height 6 ft above the top.
6. A hemispherical tank 3 ft deep in the center is full of water. Find the work done in pumping the water over the edge of the tank.

5. Total Force of Gravitational Attraction. Newton's law of gravitation states that two masses attract one another by a force F , whose magnitude is given by the formula

$$F = km_1m_2/r^2, \quad (5.1)$$

where m_1 and m_2 are the magnitudes of the masses, r is the distance between them, and k is a constant. This constant is called the gravitational constant and its value depends upon the units by which the masses

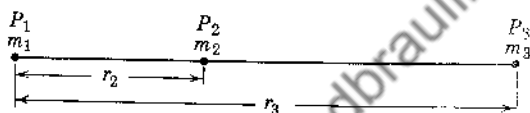


FIG. 15

and the distance are expressed. In the centimeter-gram-second system its value is $6.664 \cdot 10^{-8}$. This law assumes that the two masses are concentrated at points, which may be called *mass-points*. This is, of course, an idealization never actually reached. If the masses were not concentrated at points, it is not clear what the formula would mean, since one would not know what distance to take as r . We might expect that a summing-up process, or integration, would be necessary.

First, we shall consider three mass-points situated on a line (Fig. 15). We wish to find the amount of gravitational pull which the two masses m_2 and m_3 , situated at P_2 and P_3 , exert on m_1 at P_1 . The force exerted by m_2 is clearly

$$F_2 = km_1m_2/r_2^2,$$

and that exerted by m_3 is

$$F_3 = km_1m_3/r_3^2,$$

so that the total force of attraction which these two masses exert on m_1 is

$$F_2 + F_3 = km_1m_2/r_2^2 + km_1m_3/r_3^2.$$

If, now, instead of the two mass-points P_2 and P_3 , we imagine a mass M distributed uniformly along the line P_2P_3 , the total force will be obtained by an integration rather than by a simple addition. Let us suppose that

the length between P_2 and P_3 is L , the point P_1 is a units from P_2 and $a + L$ units from P_3 , with P_1 , P_2 , and P_3 collinear (Fig. 16). We imagine the segment P_2P_3 to be subdivided into parts, the length of a typical subdivision being $\Delta_i x$. The mass of this segment $\Delta_i x$ units long is $\delta \Delta_i x$, where δ stands for the mass per unit length, or the linear density. Since the total mass is distributed uniformly, $\delta = M/L$. Now, the mass $\delta \Delta_i x$ is not a mass-point but is distributed over a segment. If the distance

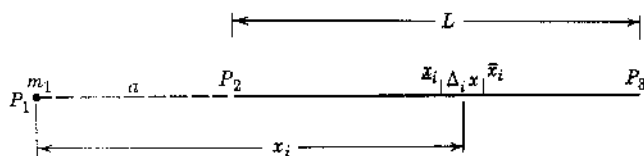


FIG. 16

from P_1 to the nearest point of the segment is x_i and that to the farthest point is $\bar{x}_i$, the force of attraction $\Delta_i F$, which the mass exerts on m_1 , lies between

$$km_1 \delta \Delta_i x / x_i^2 \quad \text{and} \quad km_1 \delta \Delta_i x / \bar{x}_i^2,$$

so that

$$\Delta_i F = km_1 \delta \Delta_i x / x_i^2, \quad \text{where} \quad x_i < x_i < \bar{x}_i.$$

The total force of attraction on m_1 is therefore the sum

$$F = \sum_{i=1}^n \Delta_i F = \sum_{i=1}^n km_1 \delta \Delta_i x / x_i^2,$$

and

$$\sum_{i=1}^n km_1 \delta \Delta_i x / x_i^2 > F > \sum_{i=1}^n km_1 \delta \Delta_i x / \bar{x}_i^2.$$

These two bounding sums must tend to a common limit as $n \rightarrow \infty$, provided that each $\Delta_i x \rightarrow 0$, and this limit we take as the value of F . Thus,

$$F = \lim_{n \rightarrow \infty} \sum_{i=1}^n km_1 \delta \Delta_i x / x_i^2 = \int_a^{a+L} (km_1 \delta / x^2) dx \\ = km_1 \delta L / a(a+L) = km_1 M / a(a+L).$$

The case just discussed enjoys a special advantage, namely, that P_1 , P_2 , and P_3 are collinear. Again, let us consider three mass-points, this time noncollinear points (Fig. 17). The magnitudes of the forces F_2 and F_3

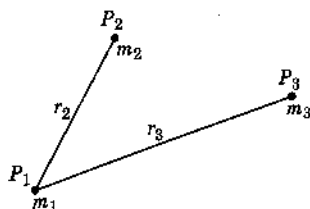


FIG. 17

are, as before, $|F_2| = km_1 m_2 / r_2^2$, $|F_3| = km_1 m_3 / r_3^2$. The absolute value sign $||$ is used here to emphasize that these formulas give us only the magnitudes of the forces in question. Their directions still have to be

considered. Formerly all directions were the same, and, hence, it was not necessary to make direction a matter of special concern. In the present case, we cannot conclude that the magnitude of the force acting on m_1 is equal to the sum of the two quantities $|F_2|$ and $|F_3|$, since the directions of these forces differ. F_2 and F_3 are vectors, and, of course, can be added by the parallelogram law, obtaining the resultant R in Fig. 18. It is more convenient, however, to obtain the components of

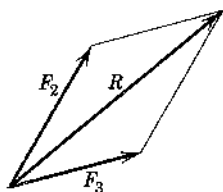


FIG. 18

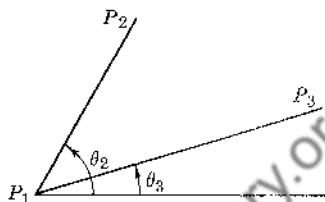


FIG. 19

F_2 and F_3 parallel to two perpendicular directions such as the x - and y -axes, and add the x -components and also the y -components to obtain the x - and y -components of the resultant. This can be done in the following way.

Let θ_2 and θ_3 be the angles on the x -axis made by lines connecting P_1 with P_2 and P_3 , respectively (Fig. 19). The x -components of F_2 and F_3 are then

$$F_2^{(x)} = (km_1m_2/r_2^2) \cos \theta_2, \quad F_3^{(x)} = (km_1m_3/r_3^2) \cos \theta_3.$$

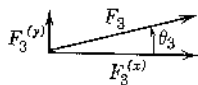
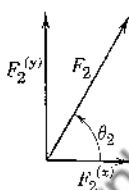


FIG. 20

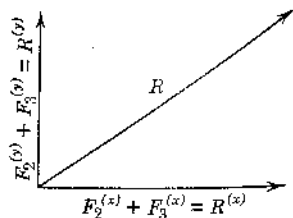


FIG. 21

Similarly, the y -components are

$$F_2^{(y)} = (km_1m_2/r_2^2) \sin \theta_2, \quad F_3^{(y)} = (km_1m_3/r_3^2) \sin \theta_3.$$

These components are shown in Fig. 20. The x -component of the resultant is now

$$R^{(x)} = F_2^{(x)} + F_3^{(x)},$$

and the y -component of the resultant is

$$R^{(y)} = F_2^{(y)} + F_3^{(y)}.$$

From these components the magnitude and direction of the resultant can be calculated (Fig. 21). We have

$$|R| = \sqrt{[R^{(x)}]^2 + [R^{(y)}]^2}$$

$$\theta = \tan^{-1} \left[\frac{R^{(y)}}{R^{(x)}} \right].$$

We should now generalize this discussion to obtain the force due to a continuous distribution of mass. This will be illustrated by the examples which follow.

EXAMPLE 1. Let a unit mass be situated at P , a units from one end of a thin homogeneous bar of length L and total mass M , measured perpendicularly from the bar (Fig. 22). Find the force by which the bar attracts this unit mass.

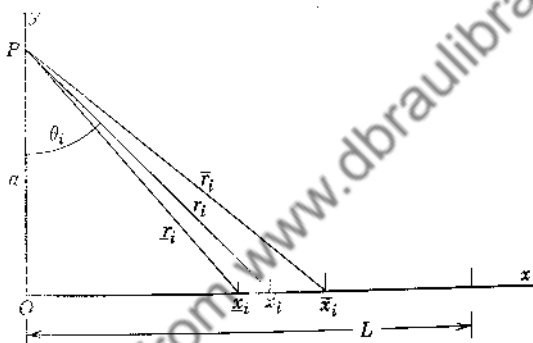


FIG. 22

SOLUTION. Let us place the coordinate axes so that P lies on the y -axis and the bar lies on the x -axis. Consider a small segment of the bar $\Delta_i x$ units long, lying between $x = x_i$ and $x = x_{i+1}$. Denoting the linear density by $\delta = M/L$, the mass of this segment is $\delta \Delta_i x$. The force of attraction $\Delta_i F$ that it exerts on the unit mass at P must lie between $k\delta \Delta_i x / r_i^2$ and $k\delta \Delta_i x / \underline{r}_i^2$, where $\bar{r}_i^2 = a^2 + \bar{x}_i^2$ and $\underline{r}_i^2 = a^2 + \underline{x}_i^2$. Hence

$$\Delta_i F = k\delta \Delta_i x / r_i^2, \quad \text{where } \bar{r}_i > r_i > \underline{r}_i$$

and

$$r_i^2 = a^2 + x_i^2, \quad \text{with } \bar{x}_i > x_i > \underline{x}_i.$$

The component of this force perpendicular to the bar, denoted by $\Delta_i F^{(y)}$ is

$$\begin{aligned} \Delta_i F^{(y)} &= (k\delta \Delta_i x / r_i^2) \cos \theta_i = (k\delta \Delta_i x / r_i^2)(a / r_i) \\ &= k\delta a \Delta_i x / (a^2 + x_i^2)^{3/2}. \end{aligned}$$

The component parallel to the bar, denoted by $\Delta_i F^{(x)}$,

$$\begin{aligned}\Delta_i F^{(x)} &= (k\delta \Delta_i x / r_i^2) \sin \theta_i = (k\delta \Delta_i x / r_i^2)(x_i / r_i) \\ &= k\delta x_i \Delta_i x / (a^2 + x_i^2)^{3/2}.\end{aligned}$$

The y -component due to the whole bar is

$$F^{(y)} = \int_0^L \frac{k\delta a}{(a^2 + x^2)^{3/2}} dx = \frac{kM}{a^2 \sqrt{a^2 + L^2}}.$$

The x -component due to the whole bar is

$$F^{(x)} = \int_0^L \frac{k\delta x}{(a^2 + x^2)^{3/2}} dx = \frac{kM}{L} \left[\frac{1}{a} - \frac{1}{\sqrt{a^2 + L^2}} \right].$$

These components determine the resultant force, just as in the case of noncollinear mass-points discussed above.

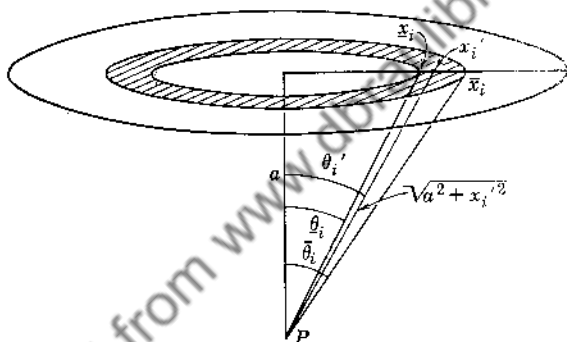


FIG. 23

EXAMPLE 2. Find the force of attraction exerted by a thin, homogeneous, circular plate with radius R , on a unit mass situated at a point P , a units from the center on a line perpendicular to the plate (Fig. 23).

SOLUTION. Consider a narrow ring concentric with the plate, with inner radius $\underline{x}_i$ and outer radius $\overline{x}_i$. The area of this ring is

$$\pi \overline{x}_i^2 - \pi \underline{x}_i^2 = \pi(\overline{x}_i + \underline{x}_i)(\overline{x}_i - \underline{x}_i).$$

Denoting the width of the ring, $\overline{x}_i - \underline{x}_i$, by $\Delta_i x$, and the average radius, $(\overline{x}_i + \underline{x}_i)/2$, by x_i , the area is equal to $\pi 2x_i \Delta_i x$. The mass of the ring is $2\pi \delta x_i \Delta_i x$, where δ = density (mass per unit area). Because of the symmetry of the ring about the axis through P , the components of attraction parallel to the plate cancel out. Hence, the force $\Delta_i F$ due to the ring is directed perpendicularly to the plate, and lies between

$$[2\pi k \delta x_i \Delta_i x / (a^2 + \underline{x}_i^2)] \cos \underline{\theta}_i \quad \text{and} \quad [2\pi k \delta x_i \Delta_i x / (a^2 + \overline{x}_i^2)] \cos \overline{\theta}_i.$$

Therefore,

$$\begin{aligned}\Delta_i F &= [2\pi k \delta x_i \Delta_i x / (a^2 + x_i'^2)] \cos \theta_i' \\ &= 2\pi k \delta a x_i \Delta_i x / (a^2 + x_i'^2)^{\frac{3}{2}},\end{aligned}$$

where x_i' lies between $\underline{x}_i$ and $\bar{x}_i$ but is not necessarily the same as x_i . By the theorem and subsequent remark in Art. 9, Chapter 7, it follows that

$$\begin{aligned}F &= \lim_{n \rightarrow \infty} \sum_{i=1}^n 2\pi k \delta a x_i \Delta_i x / (a^2 + x_i'^2)^{\frac{3}{2}} \\ &= \int_0^R \frac{2\pi k \delta a x}{(a^2 + x^2)^{\frac{3}{2}}} dx = 2\pi k \delta \left[1 - \frac{a}{\sqrt{a^2 + R^2}} \right].\end{aligned}$$

EXERCISES

1. Let a unit mass be located at a point on the perpendicular bisector of the bar in Example 1, and a units away. Find the force with which the bar attracts this unit mass.
2. Find the force with which a homogeneous rod of mass M , in the form of a semi-circle of radius a , attracts a unit mass at the center.

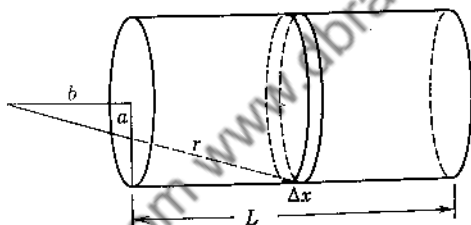


FIG. 24

3. A homogeneous cylindrical tube of radius a and length L has a total mass M . Find the force with which this mass attracts a unit mass situated on the axis of the tube, b units from one end (Fig. 24). [Hint: Sum up forces due to cylindrical rings $2\pi a \Delta x$ in area.]
4. Find the force with which a hemispherical plate attracts a unit mass at the center, if the radius is a and the plate has a density of δ units per unit area. [Hint: Sum forces due to zones of the hemisphere.]
5. Solve Exercise 4 if the unit mass is situated where the axis of the hemisphere meets the spherical surface produced.
6. Let the cylinder in Exercise 3 be solid and homogeneous. Find the attraction on the unit mass at P . [Hint: Sum forces due to cylindrical discs.]

6. Moments. Suppose a force F is applied perpendicularly to a rigid bar, as illustrated in Fig. 25. If the bar is pivoted at a point P , the force tends to produce rotation around this point. The amount of this rotational effect depends not only on F but also on the distance a , measured from P perpendicular to the line of action of F . In fact, the tendency

of F to rotate the bar is proportional to the product Fa . This product is called the *moment* of the force F about P . It is also called the *torque* due to F .

If several forces act perpendicularly on the bar, the resultant rotational effect is measured by the algebraic sum of their separate moments. In

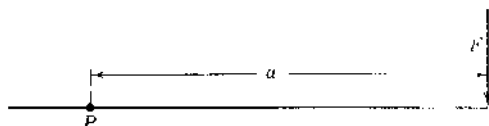


FIG. 25

Fig. 26, if distances to the right are considered positive and those to the left negative, and if upward forces are considered positive and downward forces negative, the algebraic sum of the moments is

$$F_1a_1 + F_2a_2 + F_3a_3 + F_4a_4.$$

Suppose

$$a_1 = 8, \quad a_2 = 5, \quad a_3 = 3, \quad a_4 = -4,$$

$$F_1 = -10, \quad F_2 = 15, \quad F_3 = 5, \quad F_4 = -6.$$

Then, the resultant moment is

$$(-10) \cdot 8 + 15 \cdot 5 + (-5) \cdot 3 + (-6) \cdot (-4) = 4.$$

The positive sign of this result indicates that the system of forces would rotate the bar in the counterclockwise direction, since the moment is the

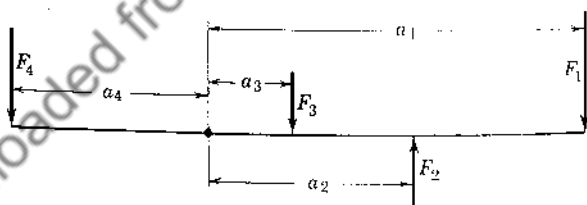


FIG. 26

same as that of a positive force acting at a positive distance, for example, a force of 1 at a distance of 4.

In many practical situations, the forces tending to rotate a body about a point are not isolated at distinct points, but are distributed continuously. A simple example is that of a bar of uniform linear density, placed on a fulcrum, as in Fig. 27. Let us suppose that the bar weighs w pounds per inch and that the distances are as indicated in the figure. In order to find the resultant moment about P we imagine the bar to be divided into n small segments having the lengths Δx . We now consider the moment of the mass of one of these segments, i.e., the moment of the force due

to the mass of this segment, $\Delta_i m = w \cdot \Delta_i x$. This moment is $w \Delta_i x \cdot x_i$, in which x_i is a distance greater than that to the near end and less than that to the far end. The resultant moment of the whole bar is the sum of all the products formed for each of the n segments. Thus, if the resultant moment is denoted by M ,

$$M = \sum_{i=1}^n x_i \Delta_i m = \sum_{i=1}^n x_i w \Delta_i x.$$

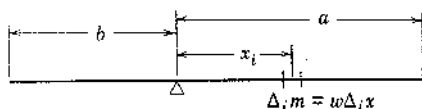


FIG. 27

Taking the limit of this sum as $n \rightarrow \infty$, we have

$$M = \int_{-b}^a w x dx = w(a^2/2 - b^2/2).$$

Let us now consider a system of forces acting at isolated points in a plane, not necessarily all on a line, each acting perpendicular to the plane.

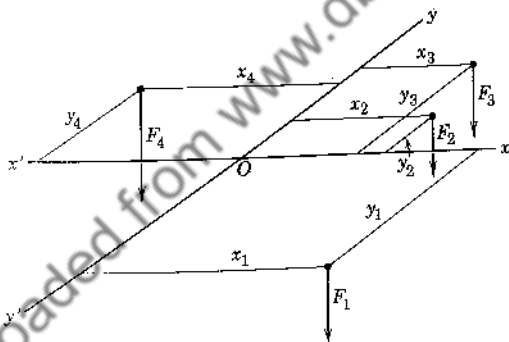


FIG. 28

Such a situation is illustrated in Fig. 28, with forces F_1, F_2, F_3 , and F_4 acting at $P_1(x_1, y_1), P_2(x_2, y_2), P_3(x_3, y_3)$, and $P_4(x_4, y_4)$, respectively. The moment of any one of these forces about an axis is defined to be the product of the force by its perpendicular distance from the axis. The moment of the whole system about the y -axis is thus

$$F_1 x_1 + F_2 x_2 + F_3 x_3 + F_4 x_4,$$

and its moment about the x -axis is

$$F_1 y_1 + F_2 y_2 + F_3 y_3 + F_4 y_4.$$

It often happens that the forces acting perpendicular to a plane are not isolated at distinct points but are distributed continuously over a portion of the plane. Let us consider a thin triangular plate OAB , situated as in Fig. 29, having a uniform density. If we mark off a narrow strip of this plate by lines parallel to the y -axis, the mass of this strip exerts downward forces, all at nearly the same distance from the y -axis. The moment of the mass of the whole strip about the y -axis is then $\Delta_i m \cdot x_i$, in which $\Delta_i m$ is the mass of the strip and x_i is a distance between the least and

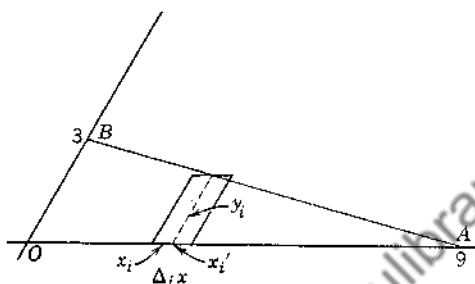


FIG. 29

greatest x in the strip. Denoting the mass per unit area by w , and the area of the strip by $\Delta_i A$, we have $\Delta_i m = w \cdot \Delta_i A$. Since the area of the strip is that of a rectangle with length y_i and width $\Delta_i x$, y_i being the average ordinate in the strip, the moment of this strip about the y -axis is $w y_i \cdot \Delta_i x \cdot x_i$. The equation of the line AB , crossing the x -axis at $x = 9$ and the y -axis at $y = 3$, is

$$x/9 + y/3 = 1, \quad \text{or} \quad y = 3 - x/3.$$

Hence,

$$y_i = 3 - x_i'/3,$$

x_i' being a value of x within the strip but not necessarily equal to x_i . Thus,

$$w y_i \cdot \Delta_i x \cdot x_i = w x_i (3 - x_i'/3) \Delta_i x.$$

As before, we conclude that the moment of the whole plate about the y -axis is

$$M = \int_0^9 w x (3 - x/3) dx = \frac{81}{2} w.$$

In the same way, we can find the moment of the plate about the x -axis. We start with a strip parallel to the x -axis, whose moment about the x -axis is

$$w x_i \Delta_i y \cdot y_i = w (9 - 3y_i') y_i \Delta_i y.$$

The moment of the whole plate about the x -axis is

$$M = \int_0^3 wy(9 - 3y) dy = \frac{27}{2} w.$$

In order to distinguish between these moments we label the moment about the y -axis M_y , and that about the x -axis M_x . The general expressions for these moments are

$$M_y = \int_a^b wlx dx, \quad (6.1)$$

$$M_x = \int_c^d wly dy, \quad (6.2)$$

in which l and l' represent the expressions for the length of a vertical strip and the length of a horizontal strip, respectively, and the limits of integration are to be determined from the area in question. In the example just given

$$l = y, l' = x, a = c = 0, b = 9, \text{ and } d = 3.$$

More generally, the moment about any line parallel to the y -axis, $x = h$, is equal to

$$\int_a^b wl(x - h) dx, \quad (6.3)$$

and the moment about any line parallel to the x -axis, $y = k$, is equal to

$$\int_c^d wly(y - k) dy. \quad (6.4)$$

The calculation of moment can be extended to solids of the types considered in Chapters 9 and 14, e.g., solids of revolution. We consider an example.

EXAMPLE 1. Extend the method of calculation of moment to find the moment about the y -axis of the solid of revolution obtained by revolving the area bounded by the parabola $x = y^2$, the x -axis, and the line $x = 4$, about the x -axis, assuming the density to be $w = \text{constant}$.

SOLUTION. We imagine a thin slice of the solid bounded by two planes perpendicular to the x -axis (Fig. 30). The moment of this slice about the y -axis is $\Delta m \cdot x = w\pi y^2 \Delta x \cdot x$, in which x lies between the greatest and least values of x in the slice, and y is the value of the ordinate that corresponds to some x in this range. We conclude that the whole moment is

$$\begin{aligned} M_y &= \int_0^4 w\pi y^2 x dx = \int_0^4 w\pi x^2 dx \\ &= \frac{64}{3}\pi w. \end{aligned}$$

Here, as before, we think of the forces due to the attraction of gravity on any part of the mass as acting perpendicular to the plane of x and y . Hence, the distance from the y -axis perpendicular to the line of the force is always x .

We shall often refer to the moment of a line, a plane, or a volume instead of a mass distributed on a line, a plane, or a volume. There are,

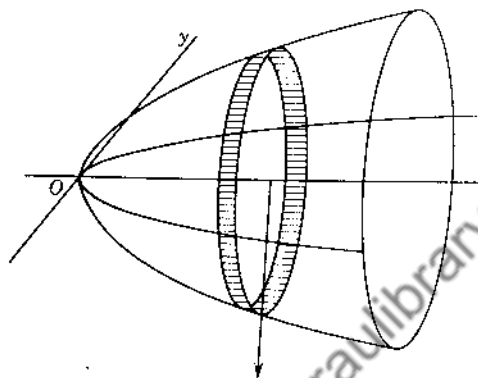


FIG. 30

of course, no actual forces associated with these moments. They are mathematical abstractions and are identical with the moments discussed above, if $w = 1$.

The moments studied in this article are a special kind, called *static moment* in mechanics. Another type of moment will be studied in Art. 8.

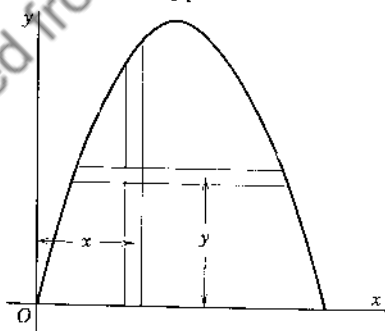


FIG. 31

EXAMPLE 2. Find the moments about the coordinate axes of the area bounded by the parabola $y = 4x - x^2$ and the x -axis (Fig. 31).

SOLUTION. The moment about the y -axis is

$$M_y = \int_0^4 xy \, dx = \int_0^4 x(4x - x^2) \, dx = \frac{64}{3}.$$

In order to find M_x we use formula 6.2, $M_x = \int_c^d l'y \, dy$, with $c = 0$, $d = 4$, and l' = distance across the parabola at height y . This distance is equal to the difference of the values of x obtained as solutions of the quadratic equation $x^2 - 4x + y = 0$. From this equation we find

$$x = 2 \pm \sqrt{4 - y}.$$

Hence,

$$l' = (2 + \sqrt{4 - y}) - (2 - \sqrt{4 - y}) = 2\sqrt{4 - y}.$$

Therefore,

$$M_x = \int_0^4 2\sqrt{4 - y} y \, dy = \frac{256}{15}.$$

EXERCISES

1. Show that the moment of a line segment about its center is 0.
2. Show that the moment of an area about an axis of symmetry is 0. [Hint: Let the axis of symmetry be the y -axis. Then $M_y = \int_{-a}^a f(x)x \, dx$, wherein $f(-x) = f(x)$. Write $M_y = \int_{-a}^0 f(x)x \, dx + \int_0^a f(x)x \, dx$ and substitute $x = -u$ in the first integral.]
3. Show that the moment of a volume of revolution about its axis is 0. [Hint: From Exercise 2 the moment of the area of a cross section is 0. Show that the moment of the volume of a cross-sectional disc is 0, and sum discs.]

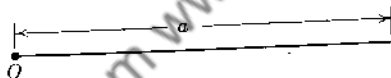


FIG. 32

4. Find the moment of the bar in Fig. 32 about the point O , if the density varies directly as the distance from O and the weight of the bar is Q pounds. [Hint:

$$Q = \int_0^a kx \, dx \text{ and } M = \int_0^a kx^2 \, dx.]$$

5. Find the moment of the area of a semicircle about the bounding diameter.
6. Find the moment of the area of a semicircle about the tangent parallel to the bounding diameter.
7. Find the moment of the area bounded by the parabola $y^2 = 4px$ and its latus rectum about the y -axis.

8. Find the moment of the area in the first quadrant bounded by the parabola $y^2 = 4px$, the latus rectum, and the x -axis, about the x -axis.

9. An isosceles triangle has its base $= b$ and altitude $= h$. (a) Find the moment of the area about the base; (b) about the line through the vertex parallel to the base.

10. A cone of revolution has the radius of its base $= r$ and altitude $= h$. (a) Find the moment of the volume about a diameter of the base; (b) about a line through the vertex parallel to the plane of the base.

11. Find the moment of a hemispherical volume about a diameter of the base.

12. The density of a hemispherical solid of radius r is equal to $1 + 1/x$, where x is the distance from the base. Find its moment about a diameter of the base.

7. Centroids. In Art. 6 we were concerned with moments resulting from the weight of a body, i.e., the action of gravity on its mass. We thought of the body as being supported at a point, or along an axis, and we calculated the moment of the weight about the point or axis. The mathematical idealization to lines, areas, and volumes identified these geometrical objects with material bodies of the same forms with density equal to unity, and hence the notion of weight, or the pull of gravity, was present, intuitively, in these cases as well as in that of a physical body. We saw in Exercises 1-3 that in certain cases the moments are 0. That is, if the bar in Exercise 1 is supported at P there is no tendency to rotate. Similarly, if any area or volume is "supported" along a line of symmetry, there is no tendency to rotate about this axis.

For a thin plate, if two lines of symmetry can be found, their intersection is a point on which the whole plate would balance. That is, there is no tendency to rotate in any direction about this point. The center of the circular plate in Fig. 33 is such a point. This point is called the *center of gravity* of the plate. For a geometrical figure, which has no mass and upon which gravity does not act, we call this point the *centroid*.

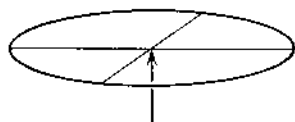


FIG. 33

DEFINITION. The centroid of a figure is a point such that, if any axis is passed through this point, the moment of the figure about this axis is 0.

Although the centroids of many figures can be observed from their symmetry, notably circles, ellipses, rectangles, and regular polygons, usually we shall have to calculate the coordinates of the centroid. For this purpose, the following theorems are useful. We state these theorems for areas. Their generalization to volumes is not difficult.

THEOREM I. The moment of an area A about the line $x = h$ is equal to $M_y - Ah$; and its moment about the line $y = k$ is equal to $M_x - Ak$.

We shall prove the first part of this theorem. From Eq. 6.1,

$$M_y = \int_a^b lx \, dx,$$

in which l is the length of a vertical strip, expressed as a function of x (see Fig. 34). From Eq. 6.3, the moment about the line $x = h$ is equal to

$$\begin{aligned} \int_a^b l(x - h) \, dx &= \int_a^b lx \, dx - \int_a^b lh \, dx \\ &= M_y - h \int_a^b l \, dx. \end{aligned}$$

But, $\int_a^b l \, dx$ is simply the area A . Hence,

$$\int_a^b l(x-h) \, dx = M_y - Ah,$$

and the first part of the theorem is proved. The second part is proved in the same way.

Theorems analogous to this one hold for moments of lines and of volumes.

THEOREM II. Let there be two axes lying in the plane of a given area and intersecting at O , about each of which the area has a moment equal to 0; then, the moment of this area about any axis through O and lying in the plane is also equal to 0.

A proof of this theorem would require some knowledge that is beyond our present scope, namely, that of double integrals. It is intuitively reasonable since, if a homogeneous plate balances on each of two intersecting axes, we would expect it to balance on their point of intersection, and hence on any axis through this point.

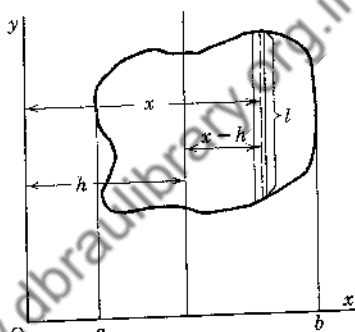


FIG. 34

The application of these theorems to the problem of finding a centroid is the following. We first use Theorem I to find values of h and k for which

$$M_y - Ah = 0; \text{ that is, } h = M_y/A \quad (7.1)$$

$$M_x - Ak = 0; \text{ that is, } k = M_x/A. \quad (7.2)$$

We denote the values of h and k satisfying Eqs. (7.1) and (7.2) by $\bar{x}$ and $\bar{y}$, respectively. The lines $x = \bar{x}$ and $y = \bar{y}$ are two intersecting axes about each of which the moment of the area is 0. By Theorem II, it follows that the moment of the area about any line through the point of intersection $(\bar{x}, \bar{y})$ is also 0. The point $(\bar{x}, \bar{y})$ is thus the centroid.

EXAMPLE 1. Find the centroid of the area in Example 2, Art. 6.

SOLUTION. Because of symmetry we know that the centroid is on the axis of the parabola; that is, $\bar{x} = 2$. In order to find $\bar{y}$, we write

$$\bar{y} = \frac{M_x}{A} = \frac{\int_0^4 2\sqrt{4-y} \, y \, dy}{\int_0^4 2\sqrt{4-y} \, dy} = \frac{256/32}{15/3} = \frac{8}{5}.$$

The coordinates of the centroid are thus $\bar{x} = 2$, $\bar{y} = \frac{8}{5}$.

EXAMPLE 2. Find the centroid of the volume of the solid in Example 1, Art. 6.

SOLUTION. Because of symmetry, we know that the centroid lies on the x -axis. In order to find the abscissa of the centroid, we use a formula analogous to (7.1). Here, $M_y = V\bar{x} = 0$; that is,

$$\bar{x} = \frac{M_y}{V} = \frac{\int_0^4 \pi y^2 x \, dx}{\int_0^4 \pi y^2 \, dx} = \frac{\int_0^4 x^2 \, dx}{\int_0^4 x \, dx} = \frac{8}{3}.$$

The centroid is on the x -axis at a distance of $\frac{8}{3}$ units from the origin.

The following alternative method of finding a moment is often convenient. We observe first that the centroid of a rectangle is its center.

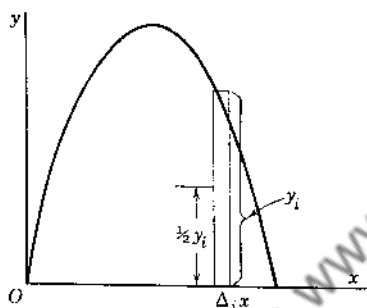


FIG. 35

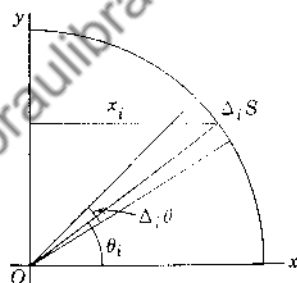


FIG. 36

It is also seen, as an immediate consequence of Theorems I and II, that the moment of a figure about any axis is equal to its content (length, area, volume) multiplied by the distance of the centroid from that axis. For example, if the centroid of an area A is $(\bar{x}, \bar{y})$, its moment about the y -axis is $A\bar{x}$. Now, consider a vertical strip of the area under the parabola $y = 4x - x^2$, approximated by a rectangle of area $y_i \Delta_i x$ (Fig. 35). The moment of this rectangle about the x -axis is $y_i \Delta_i x \cdot \frac{1}{2} y_i$, since the centroid is $\frac{1}{2} y_i$ units from the x -axis. The moment of the whole area about the x -axis is thus

$$\int_0^4 \frac{1}{2} y^2 \, dx = \int_0^4 \frac{1}{2} (4x - x^2) \, dx = \frac{256}{15}.$$

EXAMPLE 3. Find the centroid of the arc of the circle

$$x = a \cos \theta, \quad y = a \sin \theta,$$

that lies in the first quadrant, Fig. 36.

SOLUTION. Attacking this problem by the same method as that applied to the problem of finding the moment of an area, we first consider the

moment about the y -axis of a segment of the curve, $\Delta_i s$. This is $x_i \Delta_i s$. Now,*

$$\begin{aligned}\Delta_i s &\approx \sqrt{(dx/d\theta)^2 + (dy/d\theta)^2} \Delta_i \theta = \sqrt{a^2 \sin^2 \theta_i + a^2 \cos^2 \theta_i} \Delta_i \theta \\ &= a \Delta_i \theta.\end{aligned}$$

Hence, the moment of $\Delta_i s$ is approximately

$$a \Delta_i \theta \cdot x_i = a \Delta_i \theta \cdot a \cos \theta_i = a^2 \cos \theta_i \Delta_i \theta,$$

and the moment of the whole arc about the y -axis is

$$M_y = \int_0^{\pi/2} a^2 \cos \theta \, d\theta = a^2.$$

Since the length of the arc is $\pi a/2$, the moment of the arc will be equal to 0 if it is taken about the line $x = \bar{x}$, where

$$a^2 - \frac{\pi}{2} a \bar{x} = 0; \text{ that is, } \bar{x} = \frac{2a}{\pi}.$$

Since the arc is symmetrical with respect to the line $x = y$, the ordinate of the centroid is the same as its abscissa, that is,

$$\bar{y} = \frac{2a}{\pi}.$$

EXAMPLE 4. The area bounded by the curve $y = \sin x$ and the x -axis between $x = 0$ and $x = \pi$ is revolved about the y -axis. Find the centroid of the volume generated.

SOLUTION. Because of symmetry, the centroid lies on the y -axis; that is, $\bar{x} = 0$. For $\bar{y}$, we have

$$\bar{y} = M_x/V.$$

The volume is best found by the cylindrical shell method. Thus,

$$V = \int_0^{\pi} 2\pi x y \, dx = \int_0^{\pi} 2\pi x \sin x \, dx = 2\pi^2.$$

Now, the centroid of an elemental shell is its center, i.e., the point on the y -axis whose ordinate is one-half the height of the shell. The moment of the shell about the x -axis is therefore equal to its volume multiplied by one-half its height. For a typical shell this is

$$2\pi xy \Delta x \cdot \frac{1}{2} y.$$

* This is generally true as an approximation to $\Delta_i s$, as we see by reference to Eq. 7.7, Chapter 14. In the present case it can be seen at once that $\Delta_i s$ is exactly equal to $a \Delta_i \theta$.

Hence,

$$\begin{aligned} M_x &= \int_0^\pi 2\pi xy \cdot \frac{1}{2}y \, dx = \pi \int_0^\pi xy^2 \, dx \\ &= \pi \int_0^\pi x \sin^2 x \, dx \\ &= \pi^3/4. \end{aligned}$$

Therefore,

$$\bar{y} = \frac{\pi^3/4}{2\pi^2} = \frac{\pi}{8}.$$

EXERCISES

1. Find the centroid of a semicircular area.
2. Find the centroid of the area bounded by the parabola $y^2 = 4px$ and its latus rectum.
3. Find the centroid of the area in the first quadrant bounded by the parabola $y^2 = 4px$, the latus rectum, and the x -axis.
4. Prove that the centroid of a triangle is at the intersection of the medians. [Hint: Let the y -axis coincide with a median and show that $M_y = 0$.]
5. Find the centroid of a cone of revolution.
6. Find the centroid of a hemispherical solid.
7. Find the centroid of the area in the first quadrant bounded by the ellipse $x^2/a^2 + y^2/b^2 = 1$ and the coordinate axes.
8. Find the centroid of the area bounded by one arch of the curve $y = \sin x$ and the x -axis.
9. Find the centroid of the area bounded by one arch of the cycloid $x = a(\theta - \sin \theta)$, $y = a(1 - \cos \theta)$ and the x -axis.
10. Find the centroid of the area bounded by the parabola $4y = x^2$ and the line $x = y$.
11. Find the centroid of the volume generated by revolving the area in the first quadrant bounded by the ellipse $x^2/a^2 + y^2/b^2 = 1$ about the x -axis.
12. Find the centroid of the volume generated by revolving the area in Example 2, Art. 6, about the y -axis.
13. Find the centroid of the area bounded by the parabola $y = x^2$, the x -axis, and the line $x = 1$.
14. Find the centroid of the volume generated by revolving the area in Example 13 about the y -axis.
15. Find the centroid of the arc of the catenary $y = \frac{1}{2}(e^x + e^{-x})$ lying between $x = -1$ and $x = 1$.

8. Moment of Inertia. Let us consider a free mass-point (a mass concentrated at a point that is free to move in any direction). If a force F is applied to the mass, according to Newton's second law of motion it will produce an acceleration in the direction of F . The amount of the acceleration produced by a given force is inversely proportional to

the mass. In fact, force is defined as the product of the mass by the acceleration produced,

$$F = ma \quad \text{or} \quad a = F/m. \quad (8.1)$$

For example, if a mass of 2 grams is accelerated 3 cm per sec per sec, the force is $F = 2 \cdot 3 = 6$ (dynes). The *dynes* is simply the name of the unit by which force is designated in a centimeter-gram-second system. Again, if a force of 20 dynes acts on a mass-point of 2 grams, the acceleration produced is $a = 20/2 = 10$ (cm per sec per sec).

We have seen that the ability of a force to produce rotation about an axis is determined not by the force alone, but by the moment of the force, or torque, about the axis. Let us suppose that a mass-point of mass m is connected to an axis by a rigid, massless bar, r units long, Fig. 37. We ask what angular acceleration will be produced by a given torque. We have $F = ma$, where a is the acceleration in the direction of F . If α is the angular acceleration, $a = r\alpha$. Hence

$$F = mr\alpha,$$

and the torque,

$$Fr = mr^2\alpha.$$

That is,

$$\alpha = Fr/(mr^2).$$

The quantity in the denominator of this fraction is called the *moment of inertia* of m about the axis. Denoting this quantity by I , and the torque by L , we have

$$I = mr^2, \quad L = Fr, \quad \text{and} \quad \alpha = L/I. \quad (8.2)$$

Moment of inertia plays the same role in rotary motion that belongs to mass in rectilinear motion, as we see by comparison of Eqs. 8.1 and 8.2.

The moment of inertia of several mass-points is equal to the sum of their separate moments of inertia. This moment differs from that discussed in Art. 6, however, in a very important detail. There we found that in the addition of moments some might be positive and some negative, and it was quite possible for the resultant moment to be 0. The moments of inertia of two mass-points can never cancel one another, and they always add up numerically, since the quantity mr^2 is always positive, irrespective of the sign of r . Thus, the moment of inertia of the system of mass-points in Fig. 38, about the axis through 0, is $m_1r_1^2 + m_2r_2^2 + m_3r_3^2$.

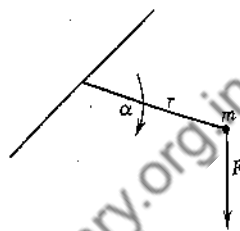


FIG. 37

The moment of inertia of a continuous distribution of mass can be obtained in a manner similar to that used in Art. 6 to obtain the moment of mass.

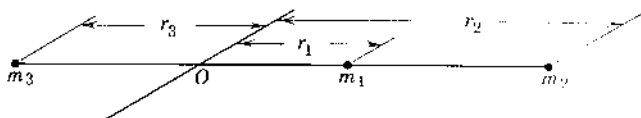


FIG. 38

EXAMPLE 1. Find the moment of inertia of a bar of linear density w and length a , about its midpoint (Fig. 39).

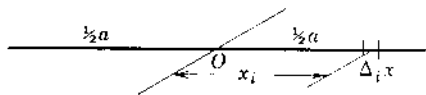


FIG. 39

SOLUTION. Let the bar be placed on the x -axis with its midpoint at 0. Imagine it to be subdivided into n segments having the lengths $\Delta_i x$ and masses $w\Delta_i x$. The moment of inertia of one of these segments about 0 must be more than $w\Delta_i x \cdot \underline{x}_i^2$ and less than $w\Delta_i x \cdot \overline{x}_i^2$, and hence equal to $w\Delta_i x \cdot x_i^2$, in which x_i is a value between the least and greatest in the segment, $\underline{x}_i$ and $\overline{x}_i$, respectively. We therefore conclude that the limit of the sum $\sum_{i=1}^n wx_i^2 \Delta_i x$ is the moment of inertia of the whole bar. Taking the limit of this sum as $n \rightarrow \infty$ and $\Delta_i x \rightarrow 0$, we have

$$I = \int_{-1/2 a}^{1/2 a} wx^2 dx = wa^3/12.$$

The moment of inertia of a thin plate can be calculated by an integral in the same way that we found the static moment of such a plate. The term *moment of inertia of an area* can be conceived as denoting the moment of inertia of a thin plate spread over the given area, with the mass per unit area equal to 1.

EXAMPLE 2. Find the moment of inertia of a circular plate about a diameter, if the radius is r and the density is constant and equal to w .

SOLUTION. Consider a thin strip parallel to the y -axis. Its area is approximately $2y\Delta x$, and its mass $w \cdot 2y \cdot \Delta x$. The moment of inertia of this strip about the y -axis is approximately $w \cdot 2y \cdot \Delta x \cdot x^2$. Reasoning as in Example 1, we conclude that

$$I = \int_{-r}^r 2wyx^2 dx.$$

Since the equation of the circle is $x^2 + y^2 = r^2$, $y = \sqrt{r^2 - x^2}$, and

$$I = \int_{-r}^r 2wx^2\sqrt{r^2 - x^2} dx$$

$$= \frac{\pi wr^4}{4}.$$

EXERCISES

1. Find the moment of inertia of the bar in Example 1, about one end.
2. Find the moment of inertia of the area of a rectangle of length a and width b about one end. About one side.
3. Find the moment of inertia of an isosceles triangle with base b and altitude h , about the base. About a line through the vertex parallel to the base.
4. Show that the results in Exercise 3 would be the same for all triangles having the same base and altitude.
5. Find the moment of inertia of the area bounded by the parabola $y^2 + 4x - 4 = 0$ and its latus rectum, about the latus rectum.
6. Find the moment of inertia of the circular plate in Example 2, about a tangent.
7. A bar 10 units long has a mass of 100, and its density varies directly as the distance from one end. Find the moment of inertia about the lighter end. About the heavier end. [Hint: See Exercise 4, Art. 6.]
8. Find the moment of inertia of the area bounded by the parabolas $y = \sqrt{x}$ and $y = x^2$ about the y -axis.

9. Polar Moment of Inertia. Volumes of Revolution. The method of Art. 8 enabled us to find the moment of inertia of a plane area about an axis lying in the same plane. In applications, it is often necessary to find the moment of inertia of an area about an axis perpendicular to the plane. Thus, Example 2, Art. 8 might be replaced by the following.

EXAMPLE 1. Find the moment of inertia of a circular plate about the axis perpendicular to its plane and through the center, if the radius is a and the area density is w .

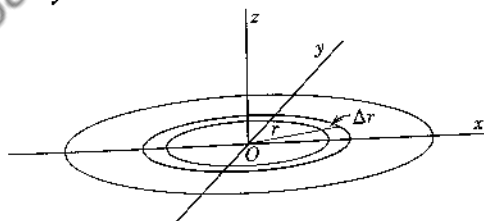


FIG. 40

SOLUTION. We consider a thin ring with O as center and width Δr (Fig. 40). The area of this ring is approximately $2\pi r \Delta r$; its mass, $w \cdot 2\pi r \cdot \Delta r$; and its moment of inertia about the axis Oz , $w \cdot 2\pi r \cdot \Delta r \cdot r^2$.

Reasoning as in Art. 8, we find the moment of inertia of the whole plate to be

$$I = \int_b^a 2\pi wr^3 dr = \frac{1}{2}\pi wa^4.$$

Since the moment of inertia of an area about a perpendicular axis can be looked upon as its moment of inertia about a point, it is natural to think of this point as the pole in a polar coordinate system. The distance of any point in the given area is now r . Moment of inertia about the pole is called *polar moment of inertia*. Methods developed up to this time are inadequate to deal with the calculation of this quantity except in the simplest cases.* The method of Example 1, however, can be extended to the calculation of the moment of inertia of a solid of revolution about its axis.

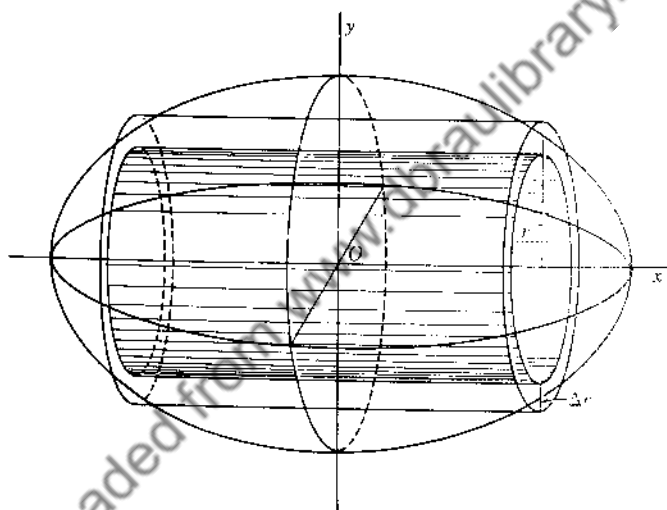


FIG. 41

EXAMPLE 2. Find the moment of inertia about its axis of an ellipsoid of revolution.

SOLUTION. Let the equation of the ellipse be

$$x^2/a^2 + y^2/b^2 = 1,$$

and suppose that it is revolved about the x -axis. Consider a cylindrical shell of radius r , thickness Δr , and length $2x$ (Fig. 41). Its volume is approximately

$$2\pi r \cdot \Delta r \cdot 2x = 4\pi x r \cdot \Delta r,$$

* This can be done in more general cases by multiple integration. See Chapter 25.

and its moment of inertia about the x -axis is approximately

$$4\pi x r \cdot \Delta r \cdot r^2 = 4\pi x r^3 \cdot \Delta r.$$

As in former problems, we conclude that the moment of inertia of the whole solid about the x -axis is

$$I = \int_0^b 4\pi x r^3 dr.$$

Here x and r are related by the equation of the ellipse, since the radius of a shell of length $2x$ is the ordinate of the point on the ellipse having the abscissa x . That is,

$$\frac{x^2}{a^2} + \frac{r^2}{b^2} = 1, \quad \text{and} \quad x = \frac{a}{b} \sqrt{b^2 - r^2}.$$

Hence,

$$\begin{aligned} I &= \int_0^b 4\pi \frac{a}{b} \sqrt{b^2 - r^2} r^3 dr \\ &= \frac{8}{15} \pi a b^4. \end{aligned}$$

EXERCISES

1. Find the moment of inertia of a cone of revolution about its axis.
2. Find the moment of inertia of a sphere about a diameter.
3. The area bounded by the parabola $y^2 = 4x$, its latus rectum, and the x -axis, is revolved about the x -axis. Find the moment of inertia of this volume about the x -axis.
4. A circular cylindrical shell has its inner radius 4 and its outer radius 6. If the height is 8, find its moment of inertia about the axis of the cylinder.
5. The area bounded by the curve $y = e^{-x}$, the axes, and the line $x = 1$, is revolved about the x -axis. Find the moment of inertia of this volume about the x -axis.

Improper Integrals

1. Extensions of the Definition of an Integral. Our use of definite integrals thus far has been limited to integrals of continuous functions over ranges of finite extent. These restrictions, if not removed, would constitute a severe handicap in the development of topics in mathematics which depend upon calculus and the extension of the uses of calculus in scientific work. Fortunately, it is easy to extend the idea of an integral to many discontinuous functions and often to ranges of infinite extent. In the present chapter we shall take up integrals of the following types:

1. $\int_a^b f(x) dx$, with $f(x)$ possessing a finite discontinuity at a or b or at some point within the interval.

EXAMPLE 1. $\int_0^2 f(x) dx$, where

$$f(x) = x \text{ for } 0 \leq x \leq 1,$$

$$= 4x - 2x^2 \text{ for } 1 < x \leq 2.$$

The graph of this function is shown in Fig. 1.

2. $\int_a^b f(x) dx$, with $f(x)$ becoming infinite at a or b or some point within the interval.

EXAMPLE 2. $\int_0^1 \frac{1}{\sqrt{1-x}} dx$. The graph is shown in Fig. 2.

3. $\int_a^\infty f(x) dx$, in which $f(x) \rightarrow 0$ as $x \rightarrow \infty$.

EXAMPLE 3. $\int_0^\infty e^{-x} dx$. The graph is shown in Fig. 3.

If one identifies the integral with the area under the graph, it is apparent that the integral in Example 1 should have a definite value, whereas this is by no means clear in Examples 2 and 3. In fact, in both examples, one's first reaction is likely to be that there can be no finite value of the integral since the figures are unlimited in one direction. It must be noted

that areas of the second and third types are bounded on one side by an asymptote. The answer to the question whether any number can be found that can reasonably be said to express such an area will depend upon how the curve approaches its asymptote. In some cases no such

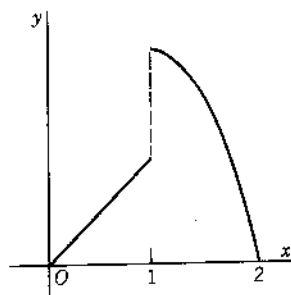


FIG. 1

number exists. Often, however, the most natural extension of the concept of area to such situations will be found to have a definite numerical value.

We shall continue to think of the area under the graph of $y = f(x)$ as the geometric interpretation of the integral, and our first problem is to define these new integrals in a way that is consistent with this interpretation.



FIG. 2

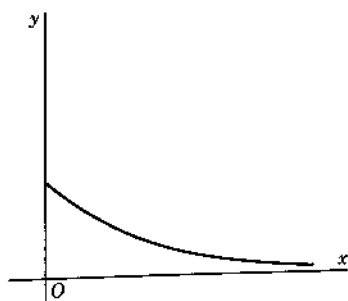


FIG. 3

2. The Integral of a Function with a Finite Discontinuity. Let us consider the area under the graph in Fig. 1. Here, no new definition of area is required. It is only necessary to agree that the area in question is that enclosed by the x -axis, the two parts of the graph of $y = f(x)$, and the dotted vertical line which connects these two parts. This area is clearly equal to

$$A = \int_0^1 x \, dx + \int_1^2 (4x - 2x^2) \, dx.$$

This result looks like an application of the theorem

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx. \quad (\text{Chapter 7, Art. 10, Theorem 3.})$$

It differs from this theorem, however, which applies to the integrals of continuous functions. For the example in question, $f(x)$ is not continuous.

Moreover, the integral $\int_1^2 (4x - 2x^2) dx$ is not the same as $\int_1^2 f(x) dx$, since $f(x)$ is not equal to $4x - 2x^2$ at $x = 1$; at this point, $f(x) = x = 1$, instead of $f(x) = 4x - 2x^2 = 2$.

If a function $f(x)$ is continuous, except at an end-point of the interval of integration, for example at b where $f(b)$ and $\lim_{x \rightarrow b} f(x)$ are finite values that may be different, we can define $\int_b^c f(x) dx$ to be equal to the integral from b to c of the continuous function created from $f(x)$ by replacing $f(b)$ by $\lim_{x \rightarrow b} f(x)$. Armed with this extension of the definition of an integral, in the above example we can write

$$\int_1^2 f(x) dx = \int_1^2 (4x - 2x^2) dx$$

and, hence,

$$A = \int_0^1 x dx + \int_1^2 (4x - 2x^2) dx = \int_0^1 f(x) dx + \int_1^2 f(x) dx.$$

Identifying the whole area with the integral over the whole interval, we write

$$\int_0^2 f(x) dx = \int_0^1 f(x) dx + \int_1^2 f(x) dx.$$

We shall formulate a definition of the integral of a function having a finite discontinuity within the interval of integration by generalizing the above example. Suppose that $f(x)$ is defined throughout the interval $a \leq x \leq c$ and is continuous except at b , within this interval; suppose also that $\lim_{h \rightarrow 0} f(b - h)$ and $\lim_{h \rightarrow 0} f(b + h)$, where $h > 0$, are two definite numbers (not infinite); then we define

$$\int_a^c f(x) dx \text{ to be } \int_a^b f(x) dx + \int_b^c f(x) dx,$$

in which

$$\int_a^b f(x) dx = \int_a^b (\text{continuous function formed by replacing } f(b) \text{ by } \lim_{h \rightarrow 0} f(b - h)) dx.$$

$$\int_b^c f(x) dx = \int_b^c (\text{continuous function formed by replacing } f(b) \text{ by } \lim_{h \rightarrow 0} f(b + h)) dx.$$

A formulation of the concept of a definite integral from a more advanced viewpoint than we have adopted permits these definitions to be proved as theorems.

A different way of obtaining an integral of a function with a finite discontinuity, which is consistent with the above definitions, will be described, because it is this method that we shall use in dealing with integrals of functions with infinite discontinuities in Art. 3.

In the above example, $f(x)$ is defined as $4x - 2x^2$ only in the interval $1 < x \leq 2$, open at the lower end. In the interval $1 + \varepsilon \leq x \leq 2$, for any ε between 0 and 1, $f(x)$ is continuous and equal to $4x - 2x^2$. Hence,

$$\int_{1+\varepsilon}^2 f(x) dx = \int_{1+\varepsilon}^2 (4x - 2x^2) dx = \frac{8}{3} - 2(1 + \varepsilon)^2 + \frac{2}{3}(1 + \varepsilon)^3,$$

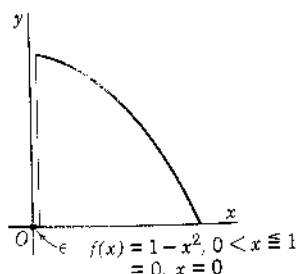


FIG. 4

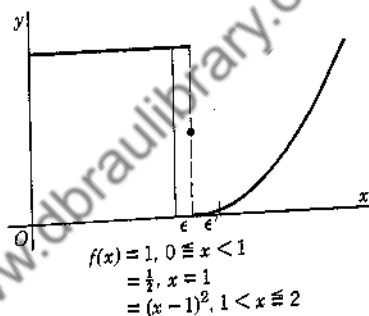


FIG. 5

As $\varepsilon \rightarrow 0$ this integral approaches $\frac{8}{3} - 2 + \frac{2}{3} = \frac{4}{3}$. Thus, if we had adopted the definition

$$\int_b^c f(x) dx = \lim_{\varepsilon \rightarrow 0} \int_{b+\varepsilon}^c f(x) dx$$

the same result as before would have been obtained. This is true in general, and we can make the following statements.

If $f(x)$ is continuous in the interval b to c , except for a finite discontinuity at one end, e.g., at b , where $\lim_{\varepsilon \rightarrow 0} f(b + \varepsilon) \neq f(b)$,

$$\int_b^c f(x) dx = \lim_{\varepsilon \rightarrow 0} \int_{b+\varepsilon}^c f(x) dx.$$

If $f(x)$ is continuous in the interval a to c , except for a finite discontinuity at a point within the interval, e.g., at b , where $\lim_{\varepsilon \rightarrow 0} f(b + \varepsilon) \neq \lim_{\varepsilon' \rightarrow 0} f(b - \varepsilon')$,

$$\int_a^c f(x) dx = \lim_{\varepsilon' \rightarrow 0} \int_a^{b-\varepsilon'} f(x) dx + \lim_{\varepsilon \rightarrow 0} \int_{b+\varepsilon}^c f(x) dx.$$

Figures 4 and 5 illustrate these statements.

For the integration of functions with finite discontinuities that are encountered in elementary calculus, the first approach given above is usually the simpler of the two.

Although we have presupposed only one discontinuity in the interval of integration, the same methods can be applied to deal with any finite number of such discontinuities.

EXAMPLE. Evaluate $\int_0^{4\pi} f(x) dx$, if

$$f(x) = \cos x, \quad 0 \leq x \leq \pi/2$$

$$= \sin x, \quad \pi/2 < x \leq \pi$$

$$= 1, \quad \pi < x < 4\pi$$

$$= 0, \quad x = 4\pi.$$



FIG. 6

SOLUTION. The graph of this function is shown in Fig. 6. Since $\lim_{h \rightarrow 0} \sin(\pi/2 + h) = \sin \pi/2$, $\lim_{h \rightarrow 0} (1 + h) = 1$, and $\lim_{h \rightarrow 0} (1 - h) = 1$, the application of the first method applied above leads to

$$\begin{aligned} \int_0^{4\pi} f(x) dx &= \int_0^{\pi/2} \cos x dx + \int_{\pi/2}^{\pi} \sin x dx + \int_{\pi}^{4\pi} 1 dx \\ &= 1 + 1 + 3\pi = 2 + 3\pi. \end{aligned}$$

EXERCISES

Calculate the following integrals.

$$1. \int_0^2 f(x) dx, \text{ where } f(x) = 1, \quad 0 \leq x < 1, \\ = 2, \quad 1 \leq x \leq 2.$$

$$2. \int_0^2 f(x) dx, \text{ where } f(x) = x^2, \quad 0 \leq x \leq 1, \\ = 4x - x^2, \quad 1 < x \leq 2.$$

$$3. \int_0^e f(x) dx, \text{ where } f(x) = e^x, \quad 0 \leq x < 1, \\ = 1, \quad x = 1, \\ = \log x/x, \quad 1 < x \leq e.$$

$$4. \int_0^1 f(x) dx, \text{ where } f(x) = 1/(1+x^2), \quad 0 \leq x < 1, \\ = 0, \quad x = 1.$$

$$5. \int_{-1}^1 f(x) dx, \text{ where } f(x) = 1/(4-x^2), \quad -1 \leq x < 0, \\ = 0, \quad x = 0, \\ = 1/(4-x^2), \quad 0 < x \leq 1.$$

3. The Integral of a Function with an Infinite Discontinuity. The extension of the integral that was discussed in Art. 2 is relatively trivial, and leads to a result that is expected. The extension to a function with an infinite discontinuity is more critical. Let us consider Example 2 of Art. 1. Since $1/\sqrt{1-x}$ becomes infinite at $x = 1$, the definition of an integral in Chapter 7 does not apply. Moreover, the interpretation of the integral as an area leads one to question whether any number can be reasonably considered to express this area, since the figure has an infinitely long spur along the asymptote $x = 1$.

It is true in this case, as in that discussed in Art. 2, that

$$\int_0^{1-\varepsilon} \frac{1}{\sqrt{1-x}} dx, \quad \varepsilon > 0,$$

exists and is readily calculated. We obtain

$$\int_0^{1-\varepsilon} \frac{1}{\sqrt{1-x}} dx = \left[-2\sqrt{1-x} \right]_0^{1-\varepsilon} = -2\sqrt{1-(1-\varepsilon)} + 2\sqrt{1} \\ = 2 - 2\sqrt{\varepsilon}.$$

This result is correct for any positive ε , no matter how small. Hence, we can say that

$$\lim_{\varepsilon \rightarrow 0} \int_0^{1-\varepsilon} \frac{1}{\sqrt{1-x}} dx = \lim_{\varepsilon \rightarrow 0} (2 - 2\sqrt{\varepsilon}) = 2.$$

The steps in this calculation show us that if we calculate the area under the curve from 0 to any ordinate drawn a little to the left of the asymptote, this area is less than 2 by a small amount $2\sqrt{\varepsilon}$, and that taking the ordinate closer and closer to the asymptote has the effect of increasing the area slightly but not enough so that it ever reaches 2. If we picture a sequence of approximating areas obtained by moving the right-hand boundary closer and closer to the asymptote, inevitably we reach a stage where we have to add enormously long strips, but by this time the widths of the strips are so minute, that the added area is inconsequential (Fig. 7).

We now define

$$\int_0^1 \frac{1}{\sqrt{1-x}} dx \text{ to be } \lim_{\epsilon \rightarrow 0} \int_0^{1-\epsilon} \frac{1}{\sqrt{1-x}} dx.$$

Hence,

$$\int_0^1 \frac{1}{\sqrt{1-x}} dx = 2.$$

Let us consider a different example, namely,

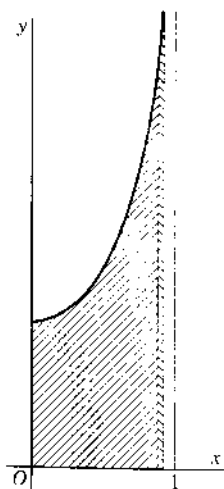


FIG. 7

$$\int_0^1 \frac{1}{(1-x)^2} dx.$$

The graph of the function $y = \frac{1}{(1-x)^2}$ in the range from 0 to 1 is similar to Fig. 7, except that the curve rises more rapidly as x approaches 1. Again the integral

$$\int_0^{1-\epsilon} \frac{1}{(1-x)^2} dx, \quad (\epsilon > 0)$$

falls under the definition in Chapter 7, since $\frac{1}{(1-x)^2}$ is continuous in the range $0 \leq x \leq 1-\epsilon$. Its value is

$$\int_0^{1-\epsilon} \frac{1}{(1-x)^2} dx = \frac{1}{\epsilon} - 1.$$

Here, if ϵ is taken smaller and smaller, $1/\epsilon - 1$ takes on larger and larger values, and $1/\epsilon - 1$ increases beyond all bounds as $\epsilon \rightarrow 0$.

From these examples, we see that the presence of an infinitely long spur of area along the asymptote may or may not make the whole area infinitely great. In the second example it did; in the first one it did not. The integrals $\int_0^1 \frac{1}{\sqrt{1-x}} dx$ and $\int_0^1 \frac{1}{(1-x)^2} dx$ are called *improper integrals*.

We say that $\int_0^1 \frac{1}{\sqrt{1-x}} dx$ *converges* to the value 2, whereas $\int_0^1 \frac{1}{(1-x)^2} dx$ does not converge, but *diverges*.

In general, if $f(x)$ is continuous for $a \leq x < b$ and $f(x) \rightarrow \pm \infty$ as $x \rightarrow b$, we define

$$\int_a^b f(x) dx \text{ to be } \lim_{\epsilon \rightarrow 0} \int_a^{b-\epsilon} f(x) dx,$$

if this limit exists. If $f(x)$ is continuous for $a < x \leq b$, while $f(x) \rightarrow \pm \infty$ as $x \rightarrow a$,

$$\int_a^b f(x) dx \text{ is } \lim_{\epsilon \rightarrow 0} \int_{a+\epsilon}^b f(x) dx,$$

if this limit exists.

If $f(x)$ becomes infinite at a point within the interval $a \leq x \leq c$, say at $x = b$, but is continuous elsewhere throughout the interval of integration, we define

$$\int_a^c f(x) dx \text{ to be } \lim_{\epsilon \rightarrow 0} \int_a^{b-\epsilon} f(x) dx + \lim_{\epsilon' \rightarrow 0} \int_{b+\epsilon'}^c f(x) dx,$$

if both of these limits exist.

We consider some further examples.

EXAMPLE 1. Evaluate $\int_0^1 \frac{1}{x} dx$, if it converges.

SOLUTION. Since $1/x$ becomes infinite as $x \rightarrow 0$, we evaluate

$$\int_{\epsilon}^1 \frac{1}{x} dx = \log 1 - \log \epsilon = -\log \epsilon.$$

As $\epsilon \rightarrow 0$, $\log \epsilon \rightarrow -\infty$. Hence, $\lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1 \frac{1}{x} dx$ does not exist, and the integral does not converge.

EXAMPLE 2. Evaluate $\int_0^1 \frac{1}{\sqrt{1-x^2}} dx$ if it converges.

SOLUTION. The function $\frac{1}{\sqrt{1-x^2}}$ becomes infinite as $x \rightarrow 1$. Hence, we evaluate

$$\int_0^{1-\epsilon} \frac{1}{\sqrt{1-x^2}} dx = \sin^{-1}(1-\epsilon).$$

As $\epsilon \rightarrow 0$, $\sin^{-1}(1-\epsilon) \rightarrow \sin^{-1} 1 = \pi/2$. Hence,

$$\int_0^1 \frac{1}{\sqrt{1-x^2}} dx = \frac{\pi}{2}.$$

EXAMPLE 3. Evaluate $\int_0^6 \frac{dx}{(x-3)^2}$ if it converges.

SOLUTION. Since $\frac{1}{(x-3)^2}$ becomes infinite at $x = 3$, we must try to calculate both of the following limits.

$$\lim_{\varepsilon \rightarrow 0} \int_0^{3-\varepsilon} \frac{dx}{(x-3)^2} \quad \text{and} \quad \lim_{\varepsilon' \rightarrow 0} \int_{3+\varepsilon'}^6 \frac{dx}{(x-3)^2}.$$

We obtain

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \int_0^{3-\varepsilon} \frac{dx}{(x-3)^2} &= \lim_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} - \frac{1}{3} \right), \\ \lim_{\varepsilon' \rightarrow 0} \int_{3+\varepsilon'}^6 \frac{dx}{(x-3)^2} &= \lim_{\varepsilon' \rightarrow 0} \left(\frac{1}{\varepsilon'} - \frac{1}{3} \right). \end{aligned}$$

Neither of these limits exists, since $1/\varepsilon$ and $1/\varepsilon'$ increase beyond all bounds.

Hence, $\int_0^6 \frac{dx}{(x-3)^2}$ does not converge.

If we were to carry out this integration formally, insert the limits, and subtract, without taking into account the fact that $1/(x-3)^2$ becomes infinite at $x = 3$, we would obtain

$$\left[-\frac{1}{x-3} \right]_0^6 = -\frac{1}{3} - \frac{1}{3} = -\frac{2}{3}.$$

This result is incorrect, since the infinite discontinuity of the integrand has been entirely neglected. That it is obviously absurd can be concluded from the fact that the integrand is positive throughout the interval and, hence, the limit of the sum $\sum_{i=1}^n f(x_i) \Delta x$ cannot be negative.

If the integrand tends to $\pm \infty$ at two points, in order for the improper integral to exist it is necessary that the integral converge at each of these points separately. For example, the integrand of $\int_0^2 \frac{2-2x}{x(2-x)} dx$ becomes infinite at $x = 0$ and $x = 2$. We cannot conclude that this integral is equal to 0, reasoning that

$$\begin{aligned} \int_0^2 \frac{2-2x}{x(2-x)} dx &= \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{2-\varepsilon} \frac{2-2x}{x(2-x)} dx \\ &= \lim_{\varepsilon \rightarrow 0} \left[\int_{\varepsilon}^{2-\varepsilon} \frac{dx}{x} - \int_{\varepsilon}^{2-\varepsilon} \frac{dx}{2-x} \right] \\ &= \lim_{\varepsilon \rightarrow 0} \left[\log \frac{2-\varepsilon}{\varepsilon} + \log \frac{\varepsilon}{2-\varepsilon} \right] = 0. \end{aligned}$$

Instead, we express the integral as the sum of two integrals, the integrand of each becoming infinite at one point only. Thus,

$$\begin{aligned}\int_0^2 \frac{2-2x}{x(2-x)} dx &= \int_0^1 \frac{2-2x}{x(2-x)} dx + \int_1^2 \frac{2-2x}{x(2-x)} dx \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^1 \frac{2-2x}{x(2-x)} dx + \lim_{\varepsilon \rightarrow 0} \int_1^{2-\varepsilon} \frac{2-2x}{x(2-x)} dx \\ &= \lim_{\varepsilon \rightarrow 0} [-\log \varepsilon(2-\varepsilon)] + \lim_{\varepsilon \rightarrow 0} [\log \varepsilon(2-\varepsilon)].\end{aligned}$$

Since neither of these limits exists, the given integral does not exist.

The intermediate point by which the interval of integration is subdivided, as the interval is here subdivided by $x = 1$, is immaterial, having no effect on the result.

EXERCISES

Evaluate the integrals that converge.

1. $\int_0^3 \frac{dx}{\sqrt{9-x^2}}$

2. $\int_0^1 \frac{dx}{1-x}$

3. $\int_3^8 \frac{dx}{x^2-9}$

4. $\int_1^9 \frac{dx}{\sqrt[3]{x-1}}$

5. $\int_1^6 \frac{9-4x}{(x-1)(6-x)} dx$

6. $\int_0^3 \frac{2x dx}{(x^2-9)^{\frac{3}{2}}}$

7. $\int_0^{\pi/2} \sec x dx$

8. $\int_0^{\pi/2} \frac{\sin x dx}{\sqrt{1-\cos x}}$

9. $\int_0^1 \frac{dx}{x^a}$ (Consider three cases. $a > 1$, $a = 1$, $a < 1$.)

10. $\int_0^4 \frac{dx}{(1-\sqrt{x})^2}$

4. An Integral over an Infinite Range. A second type of improper integral is an integral over an infinite range. This is illustrated by Example 3 of Art. 1. We obtain a satisfactory definition of this type of integral by a limiting process somewhat like that used in Art. 2. Thus, we first evaluate

$$\int_0^A e^{-x} dx = \left[-e^{-x} \right]_0^A = 1 - e^{-A} = 1 - \frac{1}{e^A}.$$

If we now let $A \rightarrow \infty$, we find that

$$\lim_{A \rightarrow \infty} \int_0^A e^{-x} dx = \lim_{A \rightarrow \infty} \left(1 - \frac{1}{e^A} \right) = 1.$$

Here we have an area that is infinitely long but which tapers off so fast that, no matter how long a segment we take, the area is never more than 1 unit. Since the limit is 1 as the length increases beyond all bounds, we say that the whole area is 1 unit. Similarly, we define the *integral*

$$\int_0^{\infty} e^{-x} dx \text{ to be } \lim_{A \rightarrow \infty} \int_0^A e^{-x} dx = 1.$$

The three kinds of unbounded ranges over which integrals may be taken are symbolized and defined as follows.

$$\int_a^{\infty} f(x) dx = \lim_{A \rightarrow \infty} \int_a^A f(x) dx, \text{ if this limit exists;}$$

$$\int_{-\infty}^a f(x) dx = \lim_{A \rightarrow -\infty} \int_A^a f(x) dx, \text{ if this limit exists;}$$

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= \int_{-\infty}^a f(x) dx + \int_a^{\infty} f(x) dx \\ &= \lim_{A' \rightarrow -\infty} \int_{A'}^a f(x) dx + \lim_{A \rightarrow \infty} \int_a^A f(x) dx, \text{ if both these limits} \end{aligned}$$

exist. In the last case, a is any number, which can be chosen in accordance with convenience.

If the limit exists in either of the first two cases, the corresponding improper integral is said to converge. It diverges if the limit does not exist. In the third case, both limits must exist in order for the integral to converge.

We shall consider some examples.

EXAMPLE 1. Evaluate $\int_0^{\infty} \frac{1}{1+x^2} dx$, if it converges.

SOLUTION. We write

$$\begin{aligned} \int_0^{\infty} \frac{1}{1+x^2} dx &= \lim_{A \rightarrow \infty} \int_0^A \frac{1}{1+x^2} dx \\ &= \lim_{A \rightarrow \infty} \tan^{-1} A = \frac{\pi}{2}. \end{aligned}$$

EXAMPLE 2. Evaluate $\int_{-\infty}^{\infty} x e^{-x^2} dx$, if it converges.

SOLUTION. For convenience, we break this infinite range into two parts— $-\infty$ to 0 and 0 to ∞ —and write

$$\begin{aligned}\int_{-\infty}^{\infty} xe^{-x^2} dx &= \int_{-\infty}^0 xe^{-x^2} dx + \int_0^{\infty} xe^{-x^2} dx \\ &= \lim_{A' \rightarrow -\infty} \int_{A'}^0 xe^{-x^2} dx + \lim_{A \rightarrow \infty} \int_0^A xe^{-x^2} dx \\ &= \lim_{A' \rightarrow -\infty} \left[-\frac{e^{-x^2}}{2} \right]_{A'}^0 + \lim_{A \rightarrow \infty} \left[-\frac{e^{-x^2}}{2} \right]_0^A \\ &= \lim_{A' \rightarrow -\infty} \left(\frac{e^{-A'^2}}{2} - \frac{1}{2} \right) + \lim_{A \rightarrow \infty} \left(\frac{1}{2} - \frac{e^{-A^2}}{2} \right) \\ &= -\frac{1}{2} + \frac{1}{2} = 0.\end{aligned}$$

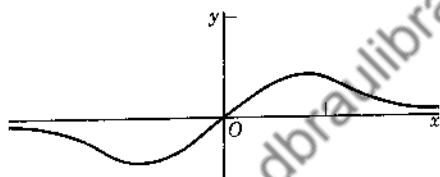


FIG. 8

Figure 8 shows the graph of $y = xe^{-x^2}$. The whole area consists of two equal parts, one above the x -axis and the other below. The above evaluation indicates that each part has an area of $\frac{1}{2}$ unit.

EXAMPLE 3. Evaluate $\int_1^{\infty} \frac{1}{\sqrt{x}} dx$, if it converges.

SOLUTION. We write

$$\int_1^{\infty} \frac{1}{\sqrt{x}} dx = \lim_{A \rightarrow \infty} \int_1^A \frac{1}{\sqrt{x}} dx = \lim_{A \rightarrow \infty} (2\sqrt{A} - 2).$$

Since $2\sqrt{A} - 2$ increases beyond all bounds as $A \rightarrow \infty$, this integral does not converge.

EXERCISES

Evaluate the integrals that converge.

1. $\int_{-\infty}^0 e^{2x} dx.$

2. $\int_3^{\infty} \frac{dx}{(x-1)^2}.$

3. $\int_1^{\infty} \frac{dx}{x-1}.$

4. $\int_3^{\infty} \frac{dx}{x^2-1}.$

$$5. \int_{-\infty}^0 \frac{dx}{(x-1)^{\frac{3}{2}}}.$$

$$6. \int_{-\infty}^{\infty} \frac{dx}{e^x + e^{-x}}.$$

$$7. \int_0^{\infty} \frac{dx}{4 + 9x^2}.$$

$$8. \int_4^{\infty} \frac{dx}{x\sqrt{x^2-4}}.$$

$$9. \int_0^{\infty} e^{-x} \sin x \, dx.$$

$$10. \int_1^{\infty} \frac{\log x}{x} \, dx.$$

$$11. \int_1^{\infty} \frac{dx}{x \log x}.$$

$$12. \int_1^a \frac{dx}{x^a} \quad (\text{Consider } a > 1, a = 1, \text{ and } a < 1.)$$

$$13. \int_1^{\infty} \frac{dx}{x^3 - 1}.$$

$$14. \int_0^{\infty} x^2 e^{-x} \, dx.$$

5. Tests for Convergence of Improper Integrals. The knowledge of the existence or nonexistence of a limit is sometimes of greater importance than the ability to calculate it if it does exist. For improper integrals, tests for convergence can be inferred from the results of Exercise 9, Art. 3, and Exercise 12, Art. 4. The tests as given below apply to positive functions, a restriction that can be removed without difficulty.

The result of Exercise 9, Art. 3, shows that the improper integral

$$\int_0^1 \frac{1}{x^a} \, dx$$

converges if $a < 1$, and diverges otherwise. The same result applies to

$$\int_0^k \frac{M}{x^a} \, dx, \quad k > 0,$$

and to

$$\int_b^k \frac{M}{(x-b)^a} \, dx, \quad k > b.$$

For example,

$$\int_2^5 \frac{8 \, dx}{(x-2)^{\frac{3}{2}}} \text{ converges.}$$

Now, if M is positive and $k > b$, and if $f(x)$ is positive and less than $M/(x-b)^a$ throughout the interval of integration,

$$\int_{b+\epsilon}^k f(x) \, dx < \int_{b+\epsilon}^k \frac{M}{(x-b)^a} \, dx.$$

Hence, if the integral on the right-hand side of this inequality converges as $\epsilon \rightarrow 0$, so does that on the left. Thus we have the following test.

TEST I. If $f(x)$ becomes infinite as $x \rightarrow b$, we conclude that $\int_b^k f(x) dx$ converges if there is a positive number M and a number $a < 1$ such that $f(x) < \frac{M}{(x-b)^a}$ throughout the interval of integration.

Since $\int_0^1 \frac{1}{x^a} dx$ diverges if $a \geq 1$, so does $\int_b^k \frac{M}{(x-b)^a} dx, k > b$. Hence, we have the following test.

TEST II. The improper integral $\int_b^k f(x) dx$ diverges if there is a positive number M such that $f(x) > \frac{M}{x-b}$ throughout the interval of integration.

The result of Exercise 12, Art. 4, shows that the improper integral $\int_1^\infty \frac{1}{x^a} dx$ converges if $a > 1$, and diverges otherwise. The same result applies to $\int_k^\infty \frac{M}{x^a} dx, k > 0$. Hence, we have the following test.

TEST III. The improper integral $\int_k^\infty f(x) dx$ converges if there is a positive number M and a number $a > 1$ such that $f(x) < M/x^a$ for all values of x greater than k .

Since $\int_k^\infty \frac{1}{x^a} dx, k > 0$, diverges if $a \leq 1$, we have the following test.

TEST IV. The improper integral $\int_k^\infty f(x) dx$ diverges if there is a positive number M such that $f(x) > M/x$ for all values of x greater than k .

In the application of these tests, any finite part of the range in which the integrand remains finite can be neglected (see Example 1).

EXAMPLE 1. Test the improper integral $\int_0^{\pi/3} \frac{\cos x}{x} dx$ for convergence.

SOLUTION. Throughout the interval of integration, $\cos x \geq \frac{1}{2} > \frac{1}{4}$. Therefore, $(\cos x)/x > \frac{1}{4}/x$. Hence, by Test II, this integral diverges.

If the interval of integration were 0 to $\pi/2$, we could still conclude that the integral diverges, since

$$\int_0^{\pi/2} \frac{\cos x}{x} dx = \int_0^{\pi/3} \frac{\cos x}{x} dx + \int_{\pi/3}^{\pi/2} \frac{\cos x}{x} dx.$$

The second integral on the right-hand side of this equation is proper and has a definite value (although we have no formula for obtaining it),

whereas we have shown that the first diverges. Hence, the integral on the left-hand side diverges.

EXAMPLE 2. Test $\int_0^1 \frac{dx}{(1+x)\sqrt{x}}$ for convergence.

SOLUTION. Throughout the interval of integration, $\frac{1}{1+x} \leq 1 < 2$. It follows that $\frac{1}{(1+x)\sqrt{x}} < \frac{2}{\sqrt{x}} = \frac{2}{x^{\frac{1}{2}}}$. Hence, putting $M = 2$, $a = \frac{1}{2} < 1$, we have $\frac{1}{(1+x)\sqrt{x}} < \frac{M}{x^a}$, and the integral converges by Test I.

EXAMPLE 3. Test $\int_1^\infty \frac{dx}{(1+x)\sqrt{x}}$ for convergence.

SOLUTION. Here $\frac{1}{1+x} < \frac{1}{x}$. It follows that

$$\frac{1}{(1+x)\sqrt{x}} < \frac{1}{x^{\frac{3}{2}}}.$$

Hence,

$$\frac{1}{(1+x)\sqrt{x}} < \frac{M}{x^a}, \text{ where } M = 1, a = \frac{3}{2} > 1.$$

Thus, by Test III, the integral converges.

EXAMPLE 4. Test $\int_2^\infty \frac{dx}{\sqrt{x-1}}$ for convergence.

SOLUTION. Since $1/\sqrt{x-1} > 1/\sqrt{x}$ for all values of $x > 2$, it follows by Test IV that this integral diverges.

EXERCISES

Test the following improper integrals for convergence or divergence.

1. $\int_0^1 \frac{1}{x + \sqrt{x}} dx.$

2. $\int_2^\infty \frac{1}{x - \sqrt{x}} dx.$

3. $\int_0^1 \frac{e^{-x}}{\sqrt{x}} dx.$

4. $\int_0^1 \frac{e^{-x}}{x^2} dx.$

5. $\int_0^1 \frac{dx}{\sqrt[3]{x^2 + x}}.$

6. $\int_0^3 \frac{dx}{\sqrt[3]{(x-1)(x-2)}}.$

7. $\int_1^\infty \frac{dx}{\sqrt{x^2 + x}}.$

8. $\int_0^\infty \frac{dx}{(x^2 + 5x + 6)^{\frac{1}{2}}}.$

$$9. \int_0^1 \frac{e^{-x}}{\sin x} dx.$$

$$11. \int_0^\infty \frac{dx}{1+x^4}$$

$$13. \int_0^{\pi/2} \csc x dx.$$

$$15. \int_0^1 \log x dx.$$

$$17. \int_0^1 \frac{\log x}{\sin x} dx.$$

$$10. \int_0^{\pi/2} \frac{dx}{\sqrt{\tan x}}.$$

$$12. \int_0^\infty \frac{x \tan^{-1} x}{1+x^2}$$

$$14. \int_0^{\pi/2} \frac{dx}{1-\cos x}$$

$$16. \int_0^1 \frac{\log x}{x} dx.$$

$$18. \int_0^1 \frac{\log x}{x^9} dx.$$

$$[\text{Hint: } \frac{\log x}{x^9} = \frac{x^{.05} \log x}{x^{.95}} < \frac{M}{x^{.95}} \text{ for small } x.]$$

$$19. \int_1^\infty x^n e^{-x} dx.$$

$$20. \text{ Show that } \int_1^\infty \frac{dx}{x \log x} \text{ diverges, by means of } \lim_{A \rightarrow \infty} \int_1^A \frac{dx}{x \log x}.$$

$$21. \text{ Show that } \int_0^1 \frac{e^x - 1}{x} dx \text{ is not an improper integral.}$$

$$22. \text{ Use the result of Exercise 21 to test } \int_0^1 \frac{x}{(e^x - 1)^2} \text{ for convergence. [Hint: } e^x - 1 < 2x \text{ for } x \text{ near } 0.]$$

23. Show that the area bounded by the x -axis, the curve $y = 1/x$, and the ordinate $x = 1$ is infinite, whereas the volume generated by revolving this area about the x -axis is finite. Find this volume.

24. Show that the area of the surface of revolution in Exercise 23 is infinite. Thus, a surface of unlimited extent is required to bound a finite volume.

25. Show that the total force of gravitational attraction that a homogeneous rod exerts on a unit mass situated on the rod cannot be found by the method of Chapter 16, Art. 5, without some modification.

Hyperbolic Functions

1. Circular Functions. The close relationship that exists between the trigonometric functions and the circle is brought out at many points in the study of those functions. An important aspect of this relationship

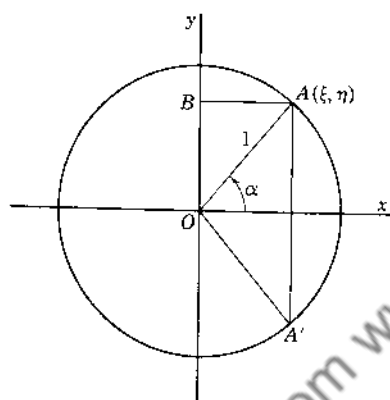


FIG. 1

lies in the way in which the trigonometric functions can be obtained by the integration of algebraic functions associated with circles. This connection was observed in Chapter 10, through such formulas as $\int \frac{dx}{\sqrt{1-x^2}} = \sin^{-1} x$, although the bearing of circles upon it was not emphasized.

Figure 1 shows the unit circle with the equation $x^2 + y^2 = 1$. Consider the sector OAA' , placed symmetrically about the x -axis. We denote the coordinates of point A by (ξ, η) (Greek xi, eta). Expressing the area S of sector OAA' in terms of the ordinate η of A , we have

$$\begin{aligned} S &= 2 \int_0^\eta \sqrt{1-y^2} dy - 2 \cdot (\text{triangle } OBA) \\ &= 2 \int_0^\eta \sqrt{1-y^2} dy - \xi\eta = 2 \int_0^\eta \sqrt{1-y^2} dy - \eta\sqrt{1-\eta^2}. \end{aligned}$$

Now, since $D_\eta(\eta\sqrt{1-\eta^2}) = (1-2\eta^2)/\sqrt{1-\eta^2}$, it follows that

$$\eta\sqrt{1-\eta^2} = \int_0^\eta \frac{1-2y^2}{\sqrt{1-y^2}} dy.$$

Hence,

$$S = \int_0^\eta \left(2\sqrt{1-y^2} - \frac{1-2y^2}{\sqrt{1-y^2}} \right) dy = \int_0^\eta \frac{1}{\sqrt{1-y^2}} dy.$$

If we take for granted the trigonometric functions and their inverses as defined in Chapter 10, we have, by formula I-8,

$$S = \int_0^{\eta} \frac{1}{\sqrt{1-y^2}} dy = \sin^{-1} \eta. \quad (1.1)$$

and

$$\eta = \sin S. \quad (1.2)$$

If, on the other hand, these functions had not previously been defined, Eqs. 1.1 and 1.2 would define them. These equations are entirely numerical, just as a definite integral is a numerical concept, in spite of the fact that we were led to them by a geometric problem. Hence, they provide strictly numerical definitions of $\sin^{-1} \eta$ and $\sin S$, that are independent of the geometrical concept of an angle and of any particular way of measuring angles. It is seen that these definitions agree with the elementary definitions since the number of radians in the angle α in Fig. 1 is the same as the number of units of area in S .

Expressing the area of sector OAA' in terms of ξ , the abscissa of A , we have

$$\begin{aligned} S &= 2 \int_{\xi}^1 \sqrt{1-x^2} dx + \text{triangle } OAA' \\ &= 2 \int_{\xi}^1 \sqrt{1-x^2} dx + \xi \eta = 2 \int_{\xi}^1 \sqrt{1-x^2} dx + \xi \sqrt{1-\xi^2}. \end{aligned}$$

As above, it can be verified by differentiation that

$$\xi \sqrt{1-\xi^2} = \int_1^{\xi} \frac{1-2x^2}{\sqrt{1-x^2}} dx.$$

Therefore,

$$\begin{aligned} S &= \int_{\xi}^1 2\sqrt{1-x^2} dx + \int_1^{\xi} \frac{1-2x^2}{\sqrt{1-x^2}} dx \\ &= -\int_1^{\xi} 2\sqrt{1-x^2} dx + \int_1^{\xi} \frac{1-2x^2}{\sqrt{1-x^2}} dx \\ &= \int_1^{\xi} \left(\frac{1-2x^2}{\sqrt{1-x^2}} - 2\sqrt{1-x^2} \right) dx = -\int_1^{\xi} \frac{1}{\sqrt{1-x^2}} dx. \end{aligned}$$

Therefore,

$$S = -\int_1^{\xi} \frac{1}{\sqrt{1-x^2}} dx = \cos^{-1} \xi, \quad (1.3)$$

and

$$\xi = \cos S. \quad (1.4)$$

The use of (1.1) and (1.3) as definitions of $\sin^{-1} S$ and $\cos^{-1} S$ is analogous to the definition of $\log x$ by an integral. All the familiar identities of trigonometry can be derived from these definitions together with (1.2) and (1.4). For example, the identity $\cos^2 S + \sin^2 S = 1$ follows from the fact that (ξ, η) lies on the unit circle $x^2 + y^2 = 1$.

EXERCISES

1. Show that the improper integral in (1.3) converges, and that the integral in (1.1) converges if $\eta \neq 1$.

2. Show that $d/dS(\sin S) = \cos S$, and $d/dS(\cos S) = -\sin S$ from Eq. 1.1 to 1.4.

2. Hyperbolic Functions. If we replace the unit circle $x^2 + y^2 = 1$ by the rectangular hyperbola $x^2 - y^2 = 1$, derivations analogous to those in Art. 1 lead to a new group of transcendental functions known as *hyperbolic*

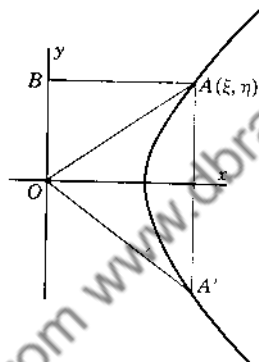


FIG. 2

functions. Again, we take a sector OAA' placed symmetrically about the x -axis (Fig. 2). Its area S , expressed in terms of the ordinate η of the point A is

$$\begin{aligned} S &= 2 \int_0^{\eta} \sqrt{y^2 + 1} \, dy - 2 \cdot (\text{triangle } OBA) \\ &= 2 \int_0^{\eta} \sqrt{y^2 + 1} \, dy - \xi\eta = 2 \int_0^{\eta} \sqrt{y^2 - 1} \, dy - \eta\sqrt{\eta^2 + 1}. \end{aligned}$$

Now,

$$D_{\eta}(\eta\sqrt{\eta^2 + 1}) = \frac{1 + 2\eta^2}{\sqrt{\eta^2 + 1}}.$$

Hence,

$$\eta\sqrt{\eta^2 + 1} = \int_0^{\eta} \frac{1 + 2y^2}{\sqrt{y^2 + 1}} \, dy,$$

and

$$S = \int_0^{\eta} \left(2\sqrt{y^2 + 1} - \frac{1 + 2y^2}{\sqrt{y^2 + 1}} \right) dy = \int_0^{\eta} \frac{1}{\sqrt{y^2 + 1}} dy.$$

Just as we defined the inverse circular function $\sin^{-1}$ by Eq. 1.1, we now define an inverse hyperbolic function $\sinh^{-1}$:

$$S = \int_0^{\eta} \frac{1}{\sqrt{y^2 + 1}} dy = \sinh^{-1} \eta. \quad (2.1)$$

Considering this function to be the inverse of a function called the *hyperbolic sine* and denoted by $\sinh$, we have

$$\eta = \sinh S. \quad (2.2)$$

Expressing the area S in terms of ξ , the abscissa of A , we have

$$\begin{aligned} S &= \text{triangle } OAA' - 2 \int_1^{\xi} \sqrt{x^2 - 1} \, dx \\ &= \xi\eta - 2 \int_1^{\xi} \sqrt{x^2 - 1} \, dx = \xi\sqrt{\xi^2 - 1} - 2 \int_1^{\xi} \sqrt{x^2 - 1} \, dx. \end{aligned}$$

Now,

$$D_{\xi}(\xi\sqrt{\xi^2 - 1}) = \frac{2\xi^2 - 1}{\sqrt{\xi^2 - 1}}.$$

Hence,

$$\xi\sqrt{\xi^2 - 1} = \int_1^{\xi} \frac{2x^2 - 1}{\sqrt{x^2 - 1}} dx,$$

and

$$S = \int_1^{\xi} \left(\frac{2x^2 - 1}{\sqrt{x^2 - 1}} - 2\sqrt{x^2 - 1} \right) dx = \int_1^{\xi} \frac{1}{\sqrt{x^2 - 1}} dx.$$

Denoting the function defined by this integral by $\cosh^{-1} \xi$, we have

$$S = \int_1^{\xi} \frac{1}{\sqrt{x^2 - 1}} dx = \cosh^{-1} \xi, \quad (2.3)$$

and

$$\xi = \cosh S. \quad (2.4)$$

This function is called the *hyperbolic cosine*.

Since the point (ξ, η) lies on the hyperbola, we have $\xi^2 - \eta^2 = 1$, whence

$$\cosh^2 S - \sinh^2 S = 1. \quad (2.5)$$

This is the fundamental identity relating $\sinh S$ and $\cosh S$, analogous to $\cos^2 S + \sin^2 S = 1$. The hyperbolic functions, $\tanh S$, $\operatorname{ctnh} S$, $\operatorname{sech} S$, and $\operatorname{csch} S$, are defined by the equations

$$\tanh S = (\sinh S)/(\cosh S), \quad \operatorname{ctnh} S = (\cosh S)/(\sinh S),$$

$$\operatorname{sech} S = 1/(\cosh S), \quad \operatorname{csch} S = 1/(\sinh S).$$

Addition formulas for the hyperbolic functions are similar in form to those for the circular functions. For example,

$$\sinh(A \pm B) = \sinh A \cosh B \pm \cosh A \sinh B \quad (2.6)$$

$$\cosh(A \pm B) = \cosh A \cosh B \pm \sinh A \sinh B \quad (2.7)$$

EXERCISES

1. Prove

$$(a) \ 1 - \tanh^2 A = \operatorname{sech}^2 A.$$

$$(b) \ \operatorname{ctnh}^2 A - 1 = \operatorname{csch}^2 A.$$

$$(c) \ \tanh(A \pm B) = \frac{\tanh A \pm \tanh B}{1 \pm \tanh A \tanh B}.$$

$$(d) \ \sinh 2A = 2 \sinh A \cosh A.$$

$$(e) \ \cosh 2A = \cosh^2 A + \sinh^2 A.$$

2. Show that $d/dS (\sinh S) = \cosh S$, and $d/dS (\cosh S) = \sinh S$, from Eqs. 2.1 to 2.4.

3. Use the results of Exercise 2 to derive

$$(a) \ d/dS (\tanh S) = \operatorname{sech}^2 S;$$

$$(b) \ d/dS (\operatorname{sech} S) = -\tanh S \operatorname{sech} S.$$

$$(c) \ d/dS (\operatorname{ctnh} S) = -\operatorname{csch}^2 S;$$

$$(d) \ d/dS (\operatorname{csch} S) = -\operatorname{ctnh} S \operatorname{csch} S.$$

3. Exponential Form. The integral by which $\sinh^{-1} \eta$ was defined in Eq. 2.1 can be evaluated by formula I-23.

$$\int_0^\eta \frac{dy}{\sqrt{y^2 + 1}} = \log(\eta + \sqrt{\eta^2 + 1}).$$

Equating this result to S , we have

$$\log(\eta + \sqrt{\eta^2 + 1}) = S, \quad \text{and} \quad \eta + \sqrt{\eta^2 + 1} = e^S.$$

Transposing and squaring,

$$\eta^2 + 1 = e^{2S} - 2\eta e^S + \eta^2,$$

whence

$$2\eta e^S = e^{2S} - 1, \quad \text{and} \quad \eta = \frac{e^S - e^{-S}}{2}.$$

Since $\eta = \sinh S$, we see that $\sinh S$ can be expressed in the exponential form:

$$\sinh S = \frac{e^S - e^{-S}}{2}. \quad (3.1)$$

Similarly, from Eq. 2.3 and formula I-24, we find that

$$\xi + \sqrt{\xi^2 - 1} = e^S, \quad \text{and} \quad \xi = \frac{e^S + e^{-S}}{2}.$$

Hence,

$$\cosh S = \frac{e^S + e^{-S}}{2}. \quad (3.2)$$

Equations 3.1 and 3.2 are often taken as the definitions of the hyperbolic sine and hyperbolic cosine, respectively.

EXERCISES

1. Derive exponential expressions for $\tanh S$, $\coth S$, $\operatorname{sech} S$, and $\operatorname{csch} S$.
2. Prove identities (2.5), (2.6), and (2.7) by means of (3.1) and (3.2).
3. Show that $d/dS (\sinh S) = \cosh S$, and $d/dS (\cosh S) = \sinh S$, from Eqs. 3.1 and 3.2.

4. Let $\tanh u = x$ and set x equal to the expression for $\tanh u$ from Exercise 1. Now solve for u , and show that

$$u = \frac{1}{2} \log (1+x)/(1-x), \text{ i.e., } \tanh^{-1} x = \frac{1}{2} \log (1+x)/(1-x).$$

This result is valid only in the interval $-1 < x < 1$. Why?

5. Use the method of Exercise 5 to show that

$$\coth^{-1} x = \frac{1}{2} \log (x+1)/(x-1) \quad \text{if } |x| > 1.$$

4. Properties of the Hyperbolic Functions. Graphs. From the equations

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2},$$

the following facts stand out.

1. $\sinh 0 = 0$. $\cosh 0 = 1$.
2. $\sinh (-x) = -\sinh x$. $\cosh (-x) = \cosh x$.

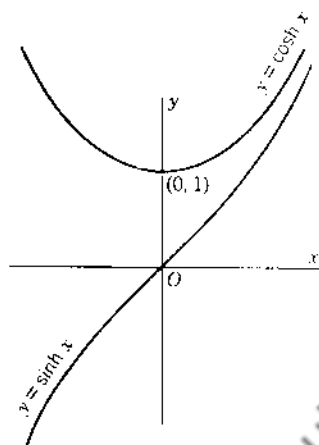
That is, $\sinh x$ is an odd function and $\cosh x$ is an even function.

3. $\sinh X \rightarrow \infty$ as $x \rightarrow \infty$. $\cosh x \rightarrow \infty$ as $x \rightarrow \infty$.
 $\rightarrow -\infty$ as $x \rightarrow -\infty$. $\rightarrow \infty$ as $x \rightarrow -\infty$.

4. Since $D(\sinh x) = \cosh x$, the slope of the curve $y = \sinh x$ is 1 at $(0, 0)$.

Since $D(\cosh x) = \sinh x$, the slope of the curve $y = \cosh x$ is 0 at $(0, 1)$. This is a minimum point, and $\cosh x \geq 1$ if $x \in \mathbb{R}$.

The graphs of these functions are shown in Fig. 3.



EXERCISES

1. Discuss the function $\tanh x$ and draw its graph.
2. Show that $\lim_{x \rightarrow 0} \frac{\sinh x}{x} = \lim_{x \rightarrow 0} \frac{e^{2x} - 1}{2x} \cdot \frac{1}{e^x} = 1$.
3. Show that $\lim_{x \rightarrow 0} \frac{\cosh x - 1}{x} = 0$, and $\lim_{x \rightarrow 0} \frac{\cosh x - 1}{x^2} = \frac{1}{2}$.

5. Integral Formulas. The hyperbolic functions furnish alternative forms for some of the formulas of Chapter 13. Thus,

$$\int \frac{dx}{\sqrt{x^2 + 1}} = \sinh^{-1} x + C.$$

This can be generalized by putting $x = u/a$ to obtain the alternative form of (I-23).

$$\int \frac{du}{\sqrt{u^2 + a^2}} = \sinh^{-1} \frac{u}{a} + C. \quad (I-23')$$

Similarly,

$$\int \frac{dx}{\sqrt{x^2 - 1}} = \cosh^{-1} x + C, \quad \text{if } x^2 \geq 1,$$

and

$$\int \frac{du}{\sqrt{u^2 - a^2}} = \cosh^{-1} \frac{u}{a} + C, \quad \text{if } u^2 \geq a^2. \quad (I-24')$$

Comparing the results of Exercises 4 and 5 of Art. 3 with formula I-22, we see that

$$\begin{aligned}\int \frac{dx}{1-x^2} &= \tanh^{-1} x + C, \quad \text{if } |x| < 1 \\ &= \operatorname{ctnh}^{-1} x + C, \quad \text{if } |x| > 1.\end{aligned}$$

Generalizing by putting $x = u/a$ in these results, we obtain

$$\int \frac{du}{a^2 - u^2} = \frac{1}{a} \tanh^{-1} \frac{u}{a} + C, \quad \text{if } |u| < |a| \quad (\text{I-22b'})$$

$$= \frac{1}{a} \operatorname{ctnh}^{-1} \frac{u}{a} + C, \quad \text{if } |u| > |a|. \quad (\text{I-22a'})$$

Formulas for integrating the hyperbolic functions are obtained from the results of such exercises as Exercises 2 and 3, Art. 2. We have

$$\int \cosh u \, du = \sinh u + C,$$

$$\int \sinh u \, du = \cosh u + C,$$

$$\int \operatorname{sech}^2 u \, du = \tanh u + C, \text{ etc. (See Exercise 1 below.)}$$

EXERCISES

1. Prove that

$$(a) \int \tanh x \, dx = \log \cosh x + C,$$

$$(b) \int \operatorname{csch}^2 x \, dx = -\operatorname{ctnh} x + C,$$

$$(c) \int \operatorname{sech} x \, dx = \tan^{-1} (\sinh x) + C,$$

$$(d) \int x \sinh x \, dx = x \cosh x - \sinh x + C.$$

2. Find the area bounded by the curve $y = \cosh x$, the x -axis, and the ordinates $x = -1$ and $x = 1$.

6. The Catenary. Imagine a perfectly flexible cable of uniform density, hung between two supports at the same height, and carrying only its own weight. The cable will fall in a curve like ACB in Fig. 4. We wish to find the equation of this curve.

Let P be a typical point on the cable, and consider the segment CP , C being the lowest point. This segment is kept in equilibrium by a horizontal force H at C , due to the pull of segment AC , and the tension T at P , due to the pull of PB . Now, T can be replaced by its vertical and horizontal components, T_v and T_h , respectively. T_v must just equal the weight of segment CP , and T_h must be equal to H , since the weight of CP cannot exert a horizontal force. Thus, T_h is the same for all points; i.e., $T_h = H = \text{constant}$.

The weight of CP is equal to its length s multiplied by its constant linear density, say w ; i.e., $T_v = ws$. Since the vector T is tangent to the curve at P , the slope there must be equal to T_v/T_h . Hence,

$$dy/dx = ws/H. \quad (6.1)$$

This is a differential equation satisfied by the equation of the curve. Since it involves three variables, x , y , and s , we change to a form in which only x and s appear, by means of the relation

$$\frac{ds}{dx} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \quad (6.2)$$

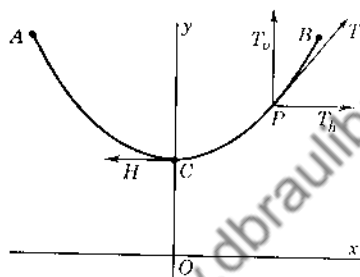


FIG. 4

Substituting ws/H for dy/dx in (6.2), we have

$$\frac{ds}{dx} = \sqrt{1 + \frac{w^2 s^2}{H^2}}, \quad \text{or} \quad \frac{dx}{ds} = \frac{1}{\sqrt{1 + s^2/a^2}}, \quad (6.3)$$

in which $1/a$ is equal to the constant w/H . From (6.3) we have

$$x = \int \frac{a ds}{\sqrt{a^2 + s^2}} = a \sinh^{-1} \frac{s}{a} + K.$$

Since $x = 0$ when $s = 0$, we have $K = 0$, and

$$x = a \sinh^{-1} \frac{s}{a}, \quad \text{or} \quad s = a \sinh \frac{x}{a}.$$

Returning to Eq. 6.1, we have

$$\frac{dy}{dx} = \frac{w}{H} a \sinh \frac{x}{a} = \sinh \frac{x}{a}, \quad \text{since} \quad \frac{w}{H} = \frac{1}{a}.$$

Therefore,

$$y = \int \sinh \frac{x}{a} dx = a \cosh \frac{x}{a} + K'.$$

Since $\cosh 0 = 1$, the constant K' is determined by the height of the cable above the x -axis. If the height of the lowest point is $a = H/w$, the value of K' is 0, and the equation of the curve is

$$y = a \cosh \frac{x}{a}, \quad \text{or} \quad y = \frac{H}{w} \cosh \frac{wx}{H}.$$

The value of H/w depends upon the ratio of the length of the cable to the distance between supports.

The form assumed by this hanging cable is a curve known as a *catenary*. In the position in which we have placed it, its equation is that given above, or the equivalent equation

$$y = a \cdot \frac{e^{x/a} + e^{-x/a}}{2}.$$

MISCELLANEOUS EXERCISES

Find $f'(x)$ in Exercises 1-10.

1. $f(x) = \cosh 4x$.

2. $f(x) = \sinh 2x \cosh 3x$.

3. $f(x) = \cosh^3 x$.

4. $f(x) = \sinh(x/2)$.

5. $f(x) = \sin x \cosh x$.

6. $f(x) = e^{\sinh x}$.

7. $f(x) = \log \sinh x$.

8. $f(x) = 2 \tan^{-1}(e^x)$.

9. $f(x) = \tan^{-1}(\sinh x)$.

10. $f(x) = \sin^{-1}(\tanh x)$.

11. Show that $\sinh x$, $\cosh x$, and $a \sinh x + b \cosh x$, each satisfy the differential equation $d^2y/dx^2 - y = 0$.

Evaluate the integrals in Exercises 12-19.

12. $\int \frac{dx}{\sqrt{1+4x^2}}$.

13. $\int \frac{dx}{\sqrt{9-4x^2}}$.

14. $\int \cosh^2 x \, dx$.

15. $\int \sinh^3 x \cosh^2 x \, dx$.

16. $\int \sinh x \log \cosh^2 x \, dx$.

17. $\int e^x \sinh x \, dx$.

18. $\int e^{-ax} \cosh bx \, dx$.

19. $\int \frac{\sinh^{-1} x}{x^2} \, dx$.

20. A body falls from rest in a medium in which the resistance is proportional to the square of the velocity. Then, $ma = mg - kv^2$, in which m is the mass, a the acceleration, g the force of gravity, v the velocity, and k is a constant. Show that

$$v = \sqrt{\frac{mg}{k}} \tanh \sqrt{\frac{kg}{m}} t.$$

Approximate Integration

1. Introduction. We know from the discussions in Chapter 7 that a definite integral of a continuous function always exists. Yet, it is clear from the introduction to Chapter 13 that such integrals cannot always be

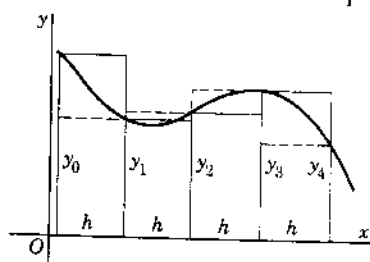


FIG. 1

found by means of inverse differentiation and the application of the fundamental theorem. In such instances, methods of approximating integrals have considerable practical value. Even some integrals that can be obtained exactly may be more conveniently approximated, owing to the difficulty of obtaining the exact form, or its unsuitability for numerical calculation. In addition to these obvious uses, approximate integration has great value in calculating areas, volumes, work, etc., when the functions to be integrated are known only graphically or through a statistical table, and are not formulated in equational form. Some of the examples at the end of this chapter illustrate such situations.

The first method of approximating the area under a continuous curve that is likely to come to mind is that which led to the formulation of the definition of the area, namely, by the sum of rectangular strips. In order to be convenient, any method of approximation should follow a formal pattern that is easily set up. Thus, we might approximate the area in Fig. 1 as follows. For convenience, we divide the interval into equal subdivisions, each denoted by h . Then, if we let the height of the rectangle approximating each strip be the length of the left-hand bounding ordinate, the area is approximated by

$$\text{Area} \approx y_0 h + y_1 h + \cdots + y_{n-1} h = h(y_0 + y_1 + \cdots + y_{n-1}). \quad (1.1)$$

If, instead, we let the height of the rectangle approximating each strip be the length of the right-hand bounding ordinate, we have

$$\text{Area} \approx y_1 h + y_2 h + \cdots + y_n h = h(y_1 + y_2 + \cdots + y_n). \quad (1.2)$$

We should expect a better approximation to be obtained by using an intermediate ordinate in each strip. A simple rule which has this effect is described in Art. 2.

2. Trapezoidal Rule. If one had no knowledge of the solution of the area problem by means of an integral, and he wished to approximate an area under a curve such as $ABQP$ in Fig. 2, he would very likely proceed as follows. He would simply replace the curve by a broken line by drawing chords, calculate the area under each chord, and add the results. For convenience in calculation, he would make the resulting trapezoidal strips of equal width. If he wanted great accuracy, he would take a large number of strips, whereas a smaller number could be used if a high degree of accuracy were unnecessary, in order to keep the labor to a minimum.

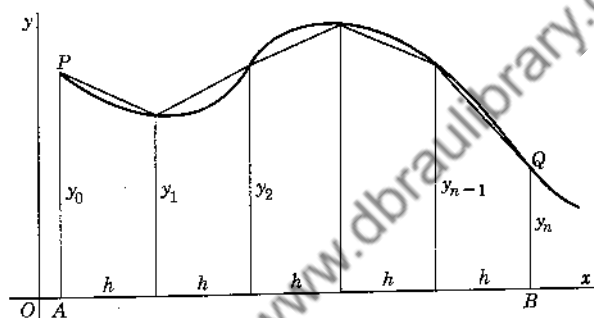


FIG. 2

Let us formalize this process. Since the area of a trapezoid is equal to one-half the sum of the parallel sides multiplied by the perpendicular distance between them, the above approximation is

$$\begin{aligned}
 \text{Area} &\approx \frac{1}{2}(y_0 + y_1)h + \frac{1}{2}(y_1 + y_2)h + \frac{1}{2}(y_2 + y_3)h + \cdots \\
 &\quad + \frac{1}{2}(y_{n-2} + y_{n-1})h + \frac{1}{2}(y_{n-1} + y_n)h \\
 &\approx \frac{1}{2}h(y_0 + y_1 + y_1 + y_2 + y_2 + y_3 + \cdots \\
 &\quad + y_{n-2} + y_{n-1} + y_{n-1} + y_n) \\
 &\approx h(\frac{1}{2}y_0 + y_1 + y_2 + \cdots + y_{n-1} + \frac{1}{2}y_n). \quad (2.1)
 \end{aligned}$$

The result of this *trapezoidal rule* for approximating the area under the curve is readily recognized as the arithmetic mean of the two approximations in Art. 1.

Identifying the integral of the function $y = f(x)$, $\int_a^b f(x) dx$, with the area under the curve, trapezoidal rule (2.1) becomes

$$\int_a^b f(x) dx \approx h[\frac{1}{2}f(x_0) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{1}{2}f(x_n)]. \quad (2.2)$$

EXAMPLE 1. Approximate the area under the curve $y = 1/(1+x^4)$ from $x = 0$ to $x = 2$, letting $n = 4$ (Fig. 3).

SOLUTION. The points of subdivision are $x_0 = 0$, $x_1 = \frac{1}{2}$, $x_2 = 1$, $x_3 = \frac{3}{2}$, $x_4 = 2$; $h = \frac{1}{2}$; and the corresponding ordinates are
 $y_0 = f(x_0) = f(0) = 1/(1+0^4) = 1$, $y_1 = f(x_1) = f(\frac{1}{2}) = 1/[1+(\frac{1}{2})^4] = .94$,
 $y_2 = f(x_2) = f(1) = 1/(1+1^4) = .5$, $y_3 = f(x_3) = f(\frac{3}{2}) = 1/[1+(\frac{3}{2})^4] = .165$,
 $y_4 = f(x_4) = f(2) = 1/(1+2^4) = .059$.

Hence, by formula 2.1,

$$\text{Area} \approx \frac{1}{2}[\frac{1}{2}(1) + .94 + .5 + .165 + \frac{1}{2}(.059)] = .767.$$

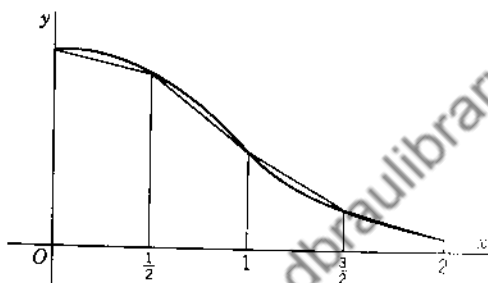


FIG. 3

EXAMPLE 2. Approximate $\int_0^3 \sqrt{1+x^3} dx$, using 6 subintervals.

SOLUTION. The points of subdivision are $x_0 = 0$, $x_1 = \frac{1}{2}$, $x_2 = 1$, $x_3 = \frac{3}{2}$, $x_4 = 2$, $x_5 = \frac{5}{2}$, $x_6 = 3$; $h = \frac{1}{2}$; and the corresponding values of $f(x)$ are

$$f(x_0) = f(0) = \sqrt{1+0^3} = 1; \quad f(x_1) = f(\frac{1}{2}) = \sqrt{1+(\frac{1}{2})^3} = 1.06;$$

$$f(x_2) = f(1) = \sqrt{1+1^3} = 1.41; \quad f(x_3) = f(\frac{3}{2}) = \sqrt{1+(\frac{3}{2})^3} = 2.09;$$

$$f(x_4) = f(2) = \sqrt{1+2^3} = 3; \quad f(x_5) = f(\frac{5}{2}) = \sqrt{1+(\frac{5}{2})^3} = 4.08;$$

$$f(x_6) = f(3) = \sqrt{1+3^3} = 5.29.$$

Hence, by formula 2.2,

$$\int_0^3 \sqrt{1+x^3} dx \approx \frac{1}{2}[\frac{1}{2}(1) + 1.06 + 1.41 + 2.09 + 3 + 4.08 + \frac{1}{2}(5.29)] = 7.39.$$

EXERCISES

Approximate each of the following integrals by means of the trapezoidal rule, using the number of subintervals indicated.

$$1. \int_1^5 \sqrt{x^2-1} dx, \quad n = 4. \quad (\text{Use a table of square roots.})$$

2. $\int_0^5 \sqrt[3]{125 - x^2} \, dx, \quad n = 5.$ (Use a table of cube roots.)

3. $\int_0^2 e^{-.5x^2} \, dx, \quad n = 4.$ (Use a table of exponentials.)

4. $\int_1^4 (\log x)^2 \, dx, \quad n = 6.$

5. $\int_6^{10} y \, dx, \quad n = 10;$ given the following values of y .

x	0	1	2	3	4	5	6	7	8	9	10
y	1.72	1.60	1.44	1.24	1.06	0.92	0.80	0.70	0.63	0.56	0.50

6. $\int_2^8 \frac{1}{f(x) - x} \, dx, \quad n = 6;$ given the following values of $f(x)$.

x	2	3	4	5	6	7	8
$f(x)$	4.50	5.02	5.58	6.12	6.90	7.68	8.62

3. Simpson's Rule. The trapezoidal rule is based on the principle of approximating the curve by a series of linear segments, and hence approximating the function to be integrated by a series of functions of the first

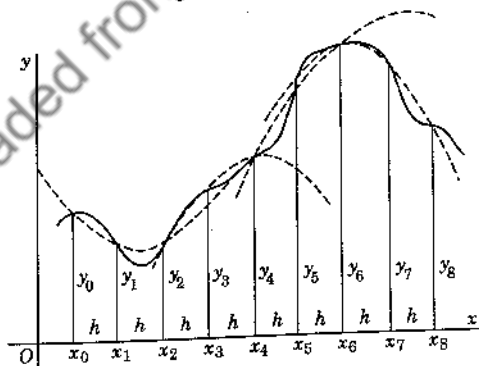


FIG. 4

degree. We now consider a method based on second-degree approximations. We imagine the curve $y = f(x)$ to be approximated by a series of parabolas, each with a vertical axis and passing through three points of the curve that are equally spaced in the direction of x (Fig. 4).

Let us consider one of these parabolas, e.g., the one determined by the points (x_0, y_0) , (x_1, y_1) , and (x_2, y_2) . Since the equation of this parabola is of the second degree in x , it can be written

$$y = a + b(x - x_1) + c(x - x_1)^2.$$

The reason for putting the equation in this form will appear in the resulting simplicity of the later steps. Since $x_0 = x_1 - h$ and $x_2 = x_1 + h$, the area under the curve between x_0 and x_2 is equal to

$$\begin{aligned} A_1 &= \int_{x_1-h}^{x_1+h} [a + b(x - x_1) + c(x - x_1)^2] dx \\ &= 2ah + \frac{2}{3}ch^3 = \frac{1}{3}h(6a + 2ch^2). \end{aligned}$$

If we can now find $6a + 2ch^2$ in terms of h and the given ordinates y_0 , y_1 , and y_2 , we shall have the area under the parabola expressed in terms of the three points of the curve which determine the parabola. For this purpose, we write

$$y_0 = a + b(x_0 - x_1) + c(x_0 - x_1)^2 = a - bh + ch^2$$

$$y_1 = a + b(x_1 - x_1) + c(x_1 - x_1)^2 = a$$

$$y_2 = a + b(x_2 - x_1) + c(x_2 - x_1)^2 = a + bh + ch^2.$$

Adding the first and third of these equations to 4 times the second, we obtain

$$y_0 + 4y_1 + y_2 = 6a + 2ch^2.$$

Hence, substituting $y_0 + 4y_1 + y_2$ for $6a + 2ch^2$ in the expression found for the area, we have

$$A_1 = \frac{1}{3}h(y_0 + 4y_1 + y_2). \quad (3.1)$$

Clearly, the area under the parabola through (x_2, y_2) , (x_3, y_3) , and (x_4, y_4) can be found in exactly the same way to be

$$A_2 = \frac{1}{3}h(y_2 + 4y_3 + y_4).$$

Similarly, the area under the parabola through (x_4, y_4) , (x_5, y_5) , (x_6, y_6) is

$$A_3 = \frac{1}{3}h(y_4 + 4y_5 + y_6),$$

and that under the parabola through (x_6, y_6) , (x_7, y_7) , and (x_8, y_8) is

$$A_4 = \frac{1}{3}h(y_6 + 4y_7 + y_8).$$

By addition, the area under the four parabolas which approximate the curve in Fig. 4 is

$$\begin{aligned} A &= A_1 + A_2 + A_3 + A_4 \\ &= \frac{1}{3}h(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + 4y_5 + 2y_6 + 4y_7 + y_8). \end{aligned}$$

The area under each parabola approximates two strips, and hence, in order to use this *parabolic rule* for approximating an area under a curve, it is necessary to subdivide the interval over which the curve is described into an even number of subintervals.

Expressed in terms of the integral of $f(x)$ this rule is

$$\int_a^b f(x) dx \approx \frac{1}{3}h[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 4f(x_{n-1}) + f(x_n)]. \quad (3.2)$$

in which n is the *even* number of equal subdivisions of the interval of integration and h is the length of each subdivision. This rule is also called *Simpson's rule*.

It should be noticed that the derivation of formula 3.1 would have been equally valid if we had started with the cubic curve through (x_0, y_0) , (x_1, y_1) , and (x_2, y_2) ,

$$y = a + b(x - x_1) + c(x - x_1)^2 + d(x - x_1)^3.$$

The student can easily carry out the steps in the same way as before, and arrive at the same formula. A somewhat surprising corollary to this result is the fact that the areas under all cubic curves which pass through the above three points are the same. It also follows that if $f(x)$ is a polynomial of degree 3 or less, Simpson's rule gives the exact value of the integral of $f(x)$.

EXAMPLE 1. Approximate $\int_0^3 \sqrt{1+x^3} dx$ by Simpson's rule, using 6 subintervals.

SOLUTION. This is the same integral that was approximated in Example 2, Art. 2. Using the same functional values, we find, by Simpson's rule:

$$\begin{aligned} \int_0^3 \sqrt{1+x^3} dx &\approx \frac{1}{3} \cdot \frac{1}{2} [1 + 4(1.06) + 2(1.41) + 4(2.09) + 2 \cdot 3 \\ &\quad + 4(4.08) + 5.29] \\ &\approx \frac{1}{6}(44.03) = 7.34. \end{aligned}$$

EXAMPLE 2. Approximate $\int_{-4}^4 e^{-x^2} dx$ by Simpson's rule, using intervals with $h = .5$.

SOLUTION. Since e^{-x^2} has the same values for $x = -a$ and $x = a$, we can set

$$\int_{-4}^4 e^{-x^2} dx = 2 \int_0^4 e^{-x^2} dx.$$

If $h = .5$, there must be 9 subdivision points at 0, .5, 1, 1.5, 2, 2.5, 3, 3.5, and 4. Using a table of exponentials, $e^{-0} = 1$, $e^{-(.5)^2} = e^{-.25} = .779$,

$e^{-1} = .368$, $e^{-(1.5)^2} = e^{-2.25} = .105$, $e^{-2^2} = e^{-4} = .018$, $e^{-(2.5)^2} = e^{-6.25} = .002$, $e^{-3^2} = e^{-9} = .000$, $e^{-(3.5)^2} = e^{-12.25} = .000$, $e^{-4^2} = e^{-16} = .000$, to 3 decimal places. Inserting these results in formula 3.2,

$$\begin{aligned}\int_0^4 e^{-x^2} dx &\approx \frac{1}{3}(.5)[1 + 4(.779) + 2(.368) + 4(.105) \\ &\quad + 2(.018) + 4(.002) + 2 \cdot 0 + 4 \cdot 0 + 0] \\ &\approx .886.\end{aligned}$$

Hence, $\int_{-\infty}^4 e^{-x^2} dx \approx 2(.886) = 1.772$, to 3 decimal places. This function approaches 0 so rapidly, as x increases, that its values for $x = 3, 3.5$, and 4 are negligible in this computation. In fact,

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \approx 1.77245.$$

EXERCISES

Solve each of the following exercises by Simpson's rule.

1. Exercise 1, Art. 2.

2. Exercise 4, Art. 2.

3. Approximate $\int_0^{\pi/2} \sqrt{\sin x} dx$, with 4 subintervals.

4. Approximate $\int_0^1 \frac{dx}{1+x}$, with 4 subintervals; with 10 subintervals. Compare the results with $\log 2 = .69315$.

5. Approximate $\int_1^{729} \frac{dx}{x} = 6 \int_1^3 \frac{dx}{x}$, with $n = 2$. Thus obtain an approximate value of $\log 729$. Obtain a second approximation by means of $3 \int_1^9 \frac{dx}{x}$, $n = 8$.

6. Find the approximate length of 1 arch of the curve $y = \sin x$, with 6 subintervals.

7. Find the approximate length of the arc of the ellipse $x = 6 \cos \phi$, $y = 4 \sin \phi$ that lies in the first quadrant, with five values of ϕ . Hence, obtain the approximate length of the whole circumference.

8. Approximate the area in the first quadrant bounded by the coordinate axes and the curve $x^3 + y^3 = 1$, with 5 ordinates.

9. During an hour's automobile drive between 9:00 and 10:00, the driver observed the speedometer readings in miles per hour as follows.

Time	9:00	9:10	9:20	9:30	9:40	9:50	10:00
Speed	32	38	29	33	42	44	38

Find the approximate mileage travelled.

10. A force that moves a body from $s = 0$ to $s = 10$ is variable, and is found to have values at various points as follows.

s	0	1	2	3	4	5	6	7	8	9	10
$f(s)$	1	1.01	1.04	1.09	1.17	1.28	1.43	1.63	1.92	2.25	2.71

Approximate the work done, by use of Simpson's rule.

11. About how much crushed rock will be required to make a roadbed 1 mile long, having the cross section shown in Fig. 5?

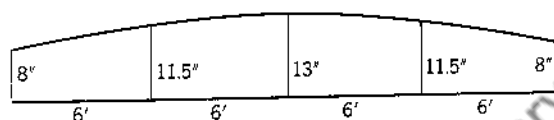


FIG. 5

12. A roadway is to be cut through a hillside. In order to estimate the amount of earth that must be removed, the area of cross sections of the cut are calculated at 10-ft intervals along the proposed roadway. The volume is then approximated by Simpson's rule applied to the formula $V = \int A(x) dx$. One of these cross sections has the form shown in Fig. 6. Approximate the area.

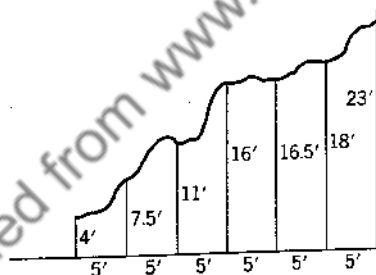


FIG. 6

13. If the areas of cross sections of the cut described in Exercise 12, at intervals of 10 ft from one end of the cut, are as shown in the table, estimate the total volume.

Distance	0	10	20	30	40	50	60	70	80
Area	0	75	95	160	300	310	215	100	0

14. Figure 7 shows an indicator diagram for a cylinder of a steam engine. Abscissas represent distances of the piston from the cylinder head. OC is the clearance, i.e., the nearest that the piston approaches the head. Steam enters the cylinder at a pressure per square inch represented by AP , the height CA corresponding to atmospheric pressure. As the piston is forced away from the head by the expanding steam

the pressure decreases as indicated by the curve PQR . The piston returns against the pressure indicated by the curve BST (atmospheric pressure from B to S). The total work done on the complete stroke is therefore the area of the diagram multiplied by the area of the piston in square inches. Find the approximate amount of work done

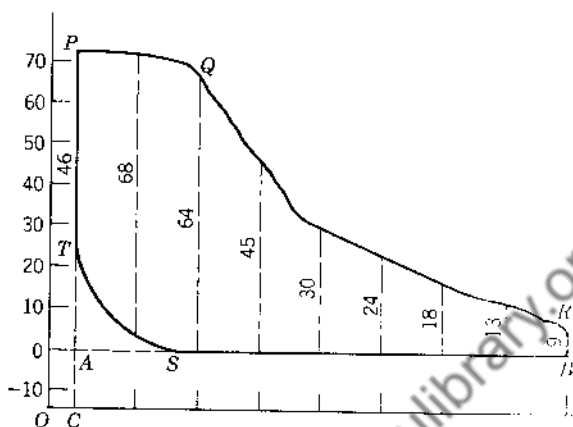


FIG. 7

on the complete stroke, using the pressure differences shown in the diagram, if the piston is 4 in. in radius and moves 12 in. forward and back.

15. A piston 10 in. in diameter has a stroke 15 in. long. Using the indicator diagram shown in Fig. 8, determine the work done on a complete stroke.

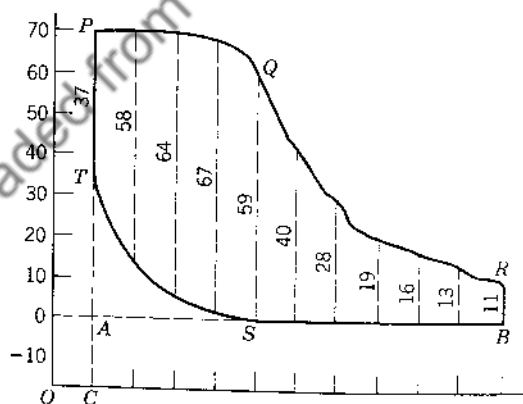


FIG. 8

4. Errors in Approximate Integrals. Workers who must rely on approximations ought to be aware of the size of the errors that may occur. We give estimates of such errors for use with trapezoidal and Simpson's rules, without proof.

Trapezoidal rule (2.2): $|\text{error}| \leq |n \overline{f''(x)} h^3/12|$, in which $|\overline{f''(x)}|$ is the maximum value of $|f''(x)|$ in the interval of integration.

Simpson's rule (3.2): $|\text{error}| \leq |n \overline{f^{(4)}(x)} h^5/180|$, in which $|\overline{f^{(4)}(x)}|$ is the maximum value of $|f^{(4)}(x)|$ in the interval of integration.*

* For further discussion of these errors see Courant, Vol. 1, p. 347. A more extended treatment is given in Milne, W. E., *Numerical Calculus*, Princeton University Press, 1949, pp. 108-122.

The Approximation of Functions by Polynomials

1. The Importance of Polynomial Approximations. Polynomials are the simplest of all functions. Only the most elementary operations are used in their construction, viz., addition, subtraction, and multiplication. On this account, not only do they afford ease in the computation of numerical values corresponding to values assumed by the independent variable, but this computation is not subject to special hazards and restrictions, as is division, in which one must be careful to guard against division by 0, root-taking which leads to ambiguous results, and the transcendental operations (e.g., finding sines or logs), in which one must either resort to limiting operations or accept approximations. Of still more importance is the fact that a polynomial is continuous and may be differentiated and integrated freely by the simplest formulas throughout the range of all real numbers.

Among the vast number of functions that do not share all the advantages of polynomials are a great many that can be approximated by polynomials with a high degree of accuracy, indeed with any desired smallness of error, obtainable by sufficiently extending the length of the polynomials involved. This fact extends to the applications of these functions many of the advantages of polynomials. For, in applications, approximations are far more common than precise values and often equally satisfactory. This chapter will be devoted to the problem of finding such polynomial approximations. Some applications of these approximations, also, will be described briefly.

2. Examples of Polynomial Approximations.

EXAMPLE 1. The fractional function $1/(1 - x)$ readily yields approximating polynomials by long division:

$$1/(1 - x) = 1 + x + x^2 + \cdots + x^{n-1} + x^n/(1 - x).$$

Thus, this fraction is approximated by $1 + x$ with error $= x^2/(1 - x)$; by $1 + x + x^2$ with error $= x^3/(1 - x)$, etc. If x is a number numerically less than 1, the error becomes less and less as more terms are included in the polynomial, since for $|x| < 1$, $|x^n|$ decreases as n increases, whereas

$1 - x$ remains fixed. If $|x| > 1$, the reverse is true, which shows that good approximations to this function are furnished by these polynomials only if $|x| < 1$. If $x = 1$, the function has no meaning nor does the expression for the error. If $x = -1$, the numerical value of the error is $\frac{1}{2}$, and does not decrease as n increases.

EXAMPLE 2.

$$1/(1+x) = 1 - x + x^2 - + \dots + (-1)^{n-1}x^{n-1} + (-1)^n x^n/(1+x).$$

This example is similar to Example 1, and similar conclusions can be drawn.

In algebra we learned a formal rule for expanding a binomial to any power. If the exponent was not a positive integer, however, this expansion was valid only for a restricted range of values. One may be tempted, nevertheless, to try to obtain polynomial approximations to such powers of a binomial by means of this expansion. Examples 1 and 2 illustrate this when the exponent is -1 . An example with the exponent -2 is

EXAMPLE 3.

$$(1-x)^{-2} = 1 + 2x + 3x^2 + \dots + (n-1)x^{n-1} + R_n(x).$$

The symbol $R_n(x)$ denotes the remainder after n terms. Here the form of the error, or remainder, is not obvious, as it was in Examples 1 and 2, and hence we cannot tell immediately how good any particular approximation is, e.g., $1 + 2x + 3x^2$. The same series of powers is obtained by differentiating Example 1.

$$1/(1-x)^2 = 1 + 2x + 3x^2 + \dots + (n-1)x^{n-1} + \left(\frac{nx^{n-1}}{1-x} + \frac{x^n}{(1-x)^2} \right)$$

Here, a formula for $R_n(x)$ is provided. As in Example 1, if $|x| < 1$, $x^n/(1-x)^2$ approaches 0 as n increases. The same is true of the term $nx^{n-1}/(1-x)$, as may be seen by reference to Chapter 11, Art. 14, Exercise 10.

EXAMPLE 4. An example of the binomial expansion for a fractional power is

$$\begin{aligned} \sqrt{1+x} &= (1+x)^{\frac{1}{2}} \\ &= 1 + \frac{1}{2}x + \frac{-\frac{1}{2} \cdot \frac{1}{2}}{2!}x^2 + \frac{-\frac{3}{2} \cdot -\frac{1}{2} \cdot \frac{1}{2}}{3!}x^3 + \dots \\ &\quad + \frac{-\frac{2n-3}{2} \cdot -\frac{2n-5}{2} \cdot \dots \cdot \frac{1}{2}}{(n-1)!}x^{n-1} + R_n(x). \end{aligned}$$

The form of $R_n(x)$ is not obvious, and we cannot judge the accuracy of polynomial approximations resulting from this expansion at this time.

From Examples 1 and 2, we can readily obtain polynomial expansions for $\log(1-x)$ and $\log(1+x)$ by integration. Thus

EXAMPLE 5. Since $\int_0^x \frac{1}{1+t} dt = \log(1+x)$ for $x > -1$, we have

$$\begin{aligned}\log(1+x) &= \int_0^x \left[1 - t + t^2 - \cdots + (-1)^{n-1} t^{n-1} + (-1)^n \frac{t^n}{1+t} \right] dt \\ &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + (-1)^{n-1} \frac{x^n}{n} + (-1)^n \int_0^x \frac{t^n}{1+t} dt.\end{aligned}$$

This result gives approximating polynomials for $\log(1+x)$ and a definite expression for the error, $R_n(x) = (-1)^n \int_0^x \frac{t^n}{1+t} dt$. We shall discuss this remainder in Art. 3.

EXAMPLE 6. By the method of Example 5 we find for $x < 1$,

$$\begin{aligned}\log(1-x) &= -\int_0^x \frac{1}{1-t} dt = -\int_0^x \left[1 - t + t^2 - \cdots + t^{n-1} - \frac{t^n}{1-t} \right] dt \\ &= -\left[x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + \frac{1}{n}x^n + \int_0^x \frac{t^n}{1-t} dt \right].\end{aligned}$$

Again we have a definite expression for the remainder, $R_n(x)$, which represents the error of the n th degree polynomial approximation.

EXAMPLE 7. The fact that $\tan^{-1} x = \int_0^x \frac{1}{1+t^2} dt$ suggests that the method of Example 5 may give us polynomial approximations to $\tan^{-1} x$. Thus, by division, as in Example 2,

$$1/(1+t^2) = 1 - t^2 + t^4 - \cdots + (-1)^{n-1} t^{2n-2} + (-1)^n \frac{t^{2n}}{1+t^2}$$

Integrating, we obtain

$$\tan^{-1} x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots + (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + (-1)^n \int_0^x \frac{t^{2n}}{1+t^2} dt.$$

The form of the remainder is again given by an integral.

3. Discussion of $R_n(x)$. An approximation, like those in Art. 2, is worthless in the absence of information as to the size of the error. Hence, a discussion of the error is an important part of every approximation problem. In particular, we need answers to the following questions.

1. Is it possible, for some range of values of x , to include enough terms in the polynomial approximation so as to make the error as small

as we please? That is, does $R_n(x)$ tend to 0 as n increases indefinitely, for the values of x in some range?

2. If the answer to question 1 is "yes," in what range is this possible? That is, for what values of x does $(R_n(x) \rightarrow 0$ as $n \rightarrow \infty$?

3. Within a specified range, or for definite values of x , what is the maximum value of the error when the approximation contains a given number of terms? That is, what bound can be set for $|R_n(x)|$ for the specified values of x ?

The relative importance of these questions depends upon the use that is to be made of the approximations. It will be seen that they enter decisively in all later discussions in this chapter and Chapter 22.

$R_n(x)$ in Examples 1, 2, and 3. Questions 1 and 2 were answered in Art. 2 for Example 1. The size of the error for any given value of x is readily computed by inserting this value in $R_n(x) = x^n/(1-x)$. Thus,

if $x = \frac{1}{2}$, the approximation $1 + x + x^2$ is in error by $R_3(\frac{1}{2}) = \frac{(\frac{1}{2})^3}{1 - \frac{1}{2}} = \frac{1}{4}$.

Since $|R_3|$ is less than this for any value of x for which $|x| < \frac{1}{2}$, we can say that, throughout the range $|x| \leq \frac{1}{2}$, $|R_3(x)| \leq \frac{1}{4}$. Although we can find the error for any value of x numerically less than 1, it is impossible to set a bound on the error resulting from 3 terms, throughout the region $|x| < 1$, since, as $x \rightarrow 1$ this error increases beyond all bounds. At the same time, if, for some value of x near 1, one finds the error of an approximating polynomial to be intolerably large, it is possible to increase the number of terms so as to reduce the error below any prescribed amount. Obviously, for values of x near 1, a good approximation would require many terms.

Similar statements can be made for Examples 2 and 3.

$R_n(x)$ in Example 5. A discussion of the error in the approximations to $\log(1+x)$ and $\log(1-x)$ offers more difficulty. Since our approximation to $\log(1+x)$ was found by integrating the expression for $1/(1+t)$, in which $R_n(t)$ decreases with increasing n if $|t| < 1$ and increases with n if $|t| > 1$, it is natural to expect the remainder in the expression for $\log(1+x)$ to behave in a like manner. Hence, we put

$$|R_n(x)| = \left| \int_0^x \frac{t^n}{1+t} dt \right| \text{ and examine this quantity for } -1 < x < 1.$$

First, let $0 \leq x < 1$. For this range, replacement of $1+t$ by 1 in the denominator of the integrand can only make the integral larger, since, in the integration $1+t$ varies from 1 to $1+x$ and 1 is its least value.

Hence, $\int_0^x \frac{t^n}{1+t} dt \leq \int_0^x t^n dt = \frac{x^{n+1}}{n+1}$, i.e., $R_n(x) \leq \frac{x^{n+1}}{n+1}$. The double sign, $\leq$, is used here only in order to include $x = 0$. Now, since $x < 1$,

the quantity x^{n+1} tends to 0 as n tends to ∞ , and so does $\frac{x^{n+1}}{n+1}$. Even for $x = 1$, the fraction $\frac{x^{n+1}}{n+1} \rightarrow 0$ as $n \rightarrow \infty$.

If $x > 1$, replacement of $1+t$ by $1+x$ in the denominator of the integrand of $R_n(x)$ has the effect of reducing the integrand. Hence, if $x > 1$,

$$\int_0^x \frac{t^n}{1+t} dt > \int_0^x \frac{t^n}{1+x} dt = \frac{1}{1+x} \int_0^x t^n dt = \frac{x^{n+1}}{(1+x)(n+1)}.$$

By the results of Art. 14, Chapter 11, if $x > 1$, the quantity $\frac{x^{n+1}}{n+1}$ increases beyond all bounds as n does, and since $1+x$ is fixed, so does $\frac{x^{n+1}}{(1+x)(n+1)}$.

If $-1 < x < 0$, we again consider the values that the denominator of the integrand assumes in the course of the integration. Here t decreases from 0 to the negative x , which is its least value, and $1+t$ decreases from 1 to $1+x$. Hence, if we replace $1+t$ by $1+x$, the integrand is increased in absolute value. Thus

$$|R_n(x)| = \left| \int_0^x \frac{t^n}{1+t} dt \right| < \left| \int_0^x \frac{t^n}{1+x} dt \right| = \left| \frac{1}{1+x} \int_0^x t^n dt \right| = \left| \frac{x^{n+1}}{(1+x)(n+1)} \right|.$$

As n tends to ∞ , $1+x$ remains fixed while x^{n+1} tends to 0. Hence, $x^{n+1}/(1+x)(n+1)$ tends to 0. Neither this expression nor the function we are trying to approximate has meaning if $x = -1$.

For $x < -1$ it can be shown, as for $x > 1$, that $R_n \rightarrow \infty$ as $n \rightarrow \infty$. Summarizing our results, we find, in answers to questions 1 and 2 at the beginning of this article: $R_n \rightarrow 0$ as $n \rightarrow \infty$ in the range $-1 < x \leq 1$. In trying to answer question 3, we cannot give a single bound for $R_n(x)$ over the entire range. In the range $0 \leq x \leq 1$, such a bound is possible.

By reference to the inequality $R_n < \frac{x^{n+1}}{n+1}$, we see that nowhere in the range $0 \leq x \leq 1$ is R_n greater than $1/(n+1)$. (See also Exercise 6, Art. 4.)

The discussion of $R_n(x)$ for Example 6 is similar to that just given. The conclusions are the same except that the range in which $R_n(x) \rightarrow 0$ as $n \rightarrow \infty$ is $-1 \leq x < 1$.

$R_n(x)$ in Example 7. Here

$$|R_n(x)| = \left| \int_0^x \frac{t^{2n}}{1+t^2} dt \right|.$$

Reasoning as in Example 5, we see that in the range $-1 \leq x \leq 1$, this integral is increased by replacing $1+t^2$ by 1 (except that if $x=0$, it is unchanged and equal to 0). Hence,

$$|R_n(x)| \leq \left| \int_0^x t^{2n} dt \right| = \left| \frac{x^{2n+1}}{2n+1} \right|.$$

Thus $|R_n(x)| \leq 1/(2n+1)$ throughout the range $|x| \leq 1$, and tends to 0 as n tends to ∞ . Outside this range $|R_n(x)| \rightarrow \infty$ as $n \rightarrow \infty$.

4. Approximation of Logarithms. The formula

$$\log(1-x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + (-1)^{n-1} \frac{1}{n}x^n + (-1)^n \int_0^x \frac{t^n}{1+t} dt \quad (4.1)$$

might be used to obtain decimal approximations to the natural logarithms. There are two reasons why this is impractical, however. First, good approximations would be possible only for numbers in the range $0 < 1+x \leq 2$, since $R_n(x)$ decreases as n increases only for x in the range $-1 < x \leq 1$. Second, except for numbers near 1, a large number of terms would be necessary to give reasonable accuracy. For example, 2000 terms would be necessary to give $\log 2$ to 3 decimal places. Extensive use of the formula

$$\log(1-x) = - \left(x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots + \frac{1}{n}x^n \right) - \int_0^x \frac{t^n}{1-t} dt \quad (4.2)$$

for this purpose is impractical for the same reason. Either of these formulas is satisfactory for a number near 1. For example, by 4.1,

$$\log 1.1 = \log(1+.1) = .1 - \frac{.01}{2} + \frac{.001}{3} - \frac{.0001}{4} + R_4$$

where $R_4 < \frac{.00001}{5} = .000002$. Thus, we find $\log 1.1 = .09531$, to 5 decimal places.

By subtracting (4.2) from (4.1), we obtain a formula that will yield the logarithms of all positive numbers. Thus,

$$\log \frac{1+x}{1-x} = 2 \left(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{2n-1}}{2n-1} \right) + \int_0^x \left[\frac{t^{2n-1}}{1+t} + \frac{t^{2n-1}}{1-t} \right] dt \quad (4.3)$$

In this formula,

$$R_n(x) = \int_0^x t^{2n-1} \left(\frac{1}{1+t} + \frac{1}{1-t} \right) dt \\ = 2 \int_0^x \frac{t^{2n-1}}{1-t^2} dt,$$

whence

$$|R_n(x)| \leq \left| \frac{x^{2n}}{n(1-x^2)} \right|, \quad \text{and} \quad |R_n| \rightarrow 0 \text{ as } n \rightarrow \infty,$$

for all values of x in the range $-1 < x < 1$.

The fractional expression $(1+x)/(1-x)$ takes on all positive values as x varies from -1 to $+1$. In order to obtain the logarithm of a large number by this formula, however, x must be taken near 1, and many terms are necessary for a close approximation. For example, $\frac{1+x}{1-x} = 20$ if $x = \frac{19}{21}$, and the polynomial in (4.3) must contain 14 terms if the error is to be less than .05. A further transformation of formula 4.3 can be found which leads to a practical formula for computing natural logarithms with a high degree of accuracy. The work is tedious at best, and since it is not an important problem of a first course in calculus, we shall leave it at this point.

EXERCISES

1. Use two terms of (4.3) to approximate $\log 1.2$, and estimate the error. [Hint: Set $(1+x)/(1-x) = 1.2$.]
2. Use two terms of (4.3) to approximate $\log 3$, and estimate the error.
3. How many terms of (4.1) are needed to approximate $\log 2$ to 1 decimal place? to 2 decimal places?
4. How many terms of (4.1) are needed to give $\log 3$ to 2 decimal places?
5. Prove the statement above that 14 terms of (4.3) are needed to approximate $\log 20$, with the error less than .05.
6. Show that for $-.9 \leq x \leq 1$, the remainder in (4.1) and (4.2) is less than $10/(n+1)$.
7. How many terms of the expression for $\tan^{-1} x$ are needed to give $\tan^{-1} .1$ with an error of less than .00001?
8. Show that if $0 < x < 1$,

$$x - x^2/2 + x^3/3(1+x) < \log(1+x) < x - x^2/2 + x^3/3.$$

Hence, find $\log 1.1$ to 4 decimal places.

5. The General Problem. In Art. 2 we obtained polynomial approximations by special methods suggested by the form of the function to be approximated or by its relationship to another function for which an approximation was already known. We seek a more general method, applicable to as wide a class of functions as possible, and independent of any previous knowledge of a relationship to known polynomials. We

first consider the preliminary question: suppose a function is known to be a polynomial of a specified degree; what further knowledge will enable us to write the polynomial?

To be explicit, let it be assumed that $f(x)$ is a polynomial in x of third degree, none of whose four coefficients is known. These coefficients can be determined by four facts about the function. For example, if we know the values assumed by the function for four different values of x , the coefficients are readily found by algebra. A different, and easier method of determining the coefficients is based on a knowledge of four facts about the function at one point, viz., the values of $f(x)$, $f'(x)$, $f''(x)$, and $f'''(x)$ for one value of x , say $x = x_1$. If $x_1 = 0$, we proceed as follows.

$$\begin{aligned} f(x) &= a + bx + cx^2 + kx^3, & f(0) &= a. \\ f'(x) &= b + 2cx + 3kx^2, & f'(0) &= b. \\ f''(x) &= 2c + 3 \cdot 2kx, & f''(0) &= 2c, \quad c = \frac{1}{2}f''(0). \\ f'''(x) &= 3 \cdot 2k, & f'''(0) &= 3 \cdot 2k, \quad k = \frac{1}{3 \cdot 2}f'''(0). \end{aligned}$$

The polynomial is, therefore,

$$f(x) = f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \frac{1}{3 \cdot 2}f'''(0)x^3. \quad (5.1)$$

EXAMPLE. Determine the polynomial of third degree, $P(x)$, if $P(0) = 3$, $P'(0) = 1$, $P''(0) = 6$, $P'''(0) = 12$.

SOLUTION. Substituting in (5.1),

$$P(x) = 3 + x + 3x^2 + 2x^3.$$

It is easy to show, in exactly the same way, that if $f(x)$ is a polynomial of n th degree, it can be written

$$f(x) = f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \cdots + \frac{1}{n!}f^{(n)}(0)x^n. \quad (5.2)$$

In a similar way, we can express a function $f(x)$, known to be a polynomial, in powers of $x-a$, in which a is any constant. Thus, if the degree of $f(x)$ is 3,

$$\begin{aligned} f(x) &= A + B(x-a) + C(x-a)^2 \\ &\quad + D(x-a)^3, & f(a) &= A \\ f'(x) &= B + 2C(x-a) + 3D(x-a)^2, & f'(a) &= B \\ f''(x) &= 2C + 3 \cdot 2D(x-a), & f''(a) &= 2C, \quad C = \frac{1}{2}f''(a) \\ f'''(x) &= 3 \cdot 2D, & f'''(a) &= 3 \cdot 2D, \quad D = \frac{1}{3 \cdot 2}f'''(a) \end{aligned}$$

and

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{3 \cdot 2}f'''(a)(x-a)^3.$$

In general, if the degree of $f(x)$ is n ,

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n. \quad (5.3)$$

Clearly, (5.2) is a special case of the general form (5.3), with $a = 0$.

If $f(x)$ is not a polynomial, it cannot be expressed exactly in a form like (5.3). This form, however, suggests a possible way of approximating more general functions by means of polynomials. If a function $f(x)$ possesses derivatives at $x = 0$, of all orders up to the $(n-1)$ st, we can write

$$f(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 + \frac{1}{3!}f'''(0)x^3 + \cdots + \frac{1}{n!}f^{(n)}(0)x^n + R_n. \quad (5.4)$$

In writing this equation, we only state that $f(x)$ is equal to a certain polynomial plus something, viz., R_n . There can be no question of our right to do this, the quantity R_n being defined (not derived nor calculated) by the equation, as the amount by which the polynomial fails to equal $f(x)$. This quantity is the error in the approximation, which may be small or large. It is a function of x and we shall henceforth write it $R_n(x)$.

If, now, for some given function $f(x)$, possessing derivatives of all orders at $x = 0$, we could show that $R_n(x) \rightarrow 0$ as $n \rightarrow \infty$, for all values of x in some interval, this would show that, by taking n sufficiently great, the error in the approximation of $f(x)$ by the polynomial consisting of the 1st n terms of (5.4) can be made as small as desired; i.e., (5.4) does in fact, furnish satisfactorily close approximations.

More generally, if $f(x)$ possesses derivatives, at $x = a$, of all orders to the $(n-1)$ st, we can write

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n + R_n(x). \quad (5.5)$$

The above remarks concerning the possibility of (5.4) furnishing satisfactorily close approximations are of equal weight for (5.5).

Functions that satisfy the conditions justifying these results have been illustrated by the examples in Art. 2. It is easily verified, for example, that the coefficients in the expression for $\log(1+x)$, in Art. 2, are identical with those given by (5.4), when $f(x) = \log(1+x)$. Moreover,

it was shown in Art. 3 that the remainder in this expression tends to 0 as n tends to ∞ , for values of x in the interval $-1 < x \leq 1$. This was made possible by the knowledge of the form of $R_n(x)$. In Art. 6 we shall find a form assumed by $R_n(x)$ in the general case.

6. Taylor's Theorem. Equations 5.4 and 5.5 give no information about a function $f(x)$, in the absence of further knowledge about the remainder $R_n(x)$. A knowledge of an expression for this remainder, from which we could calculate it, or, although lacking a precise evaluation, set bounds on its size, would constitute a very valuable item of information, as we have seen in the examples in Art. 3. It is this knowledge that is given by Taylor's theorem.* This theorem can be stated as follows.

TAYLOR'S THEOREM. Let the function $f(x)$ be continuous and possess continuous derivatives of all orders up to and including the n th, at $x = a$; then, in some interval including $x = a$, $f(x)$ is equal to a polynomial in $(x - a)$ of degree $(n - 1)$ plus a remainder $R_n(x)$ that can be put in a specified form as follows.

$$\left. \begin{aligned} f(x) = f(a) + f'(a)(x - a) + \frac{1}{2!} f''(a)(x - a)^2 + \cdots \\ + \frac{1}{(n - 1)!} f^{(n-1)}(a)(x - a)^{n-1} + R_n(x), \end{aligned} \right\} \quad (6.1)$$

in which

$$R_n(x) = \frac{1}{n!} f^{(n)}(x_1)(x - a)^n,$$

wherein x_1 lies between a and x .

Other forms of $R_n(x)$ can be found, which are more suitable for numerical calculation than the one given here. The form of $R_n(x)$ in Example 5, Art. 2, furnishes an example. For further discussion of this point, see Courant: Vol. I, pp. 322-4.

Equations 6.1 can be written in the alternative form

$$\left. \begin{aligned} f(a + h) = f(a) + f'(a)h + \frac{1}{2!} f''(a)h^2 + \cdots \\ + \frac{1}{(n - 1)!} f^{(n-1)}(a)h^{n-1} + R_n(h), \end{aligned} \right\} \quad (6.2)$$

with

$$R_n(h) = \frac{1}{n!} f^{(n)}(a + \theta h)h^n, \quad 0 < \theta < 1.$$

* Discovered by Brook Taylor (1685-1731).

The mean value theorem of differential calculus is clearly a special case of Taylor's theorem

$$f(a+h) = f(a) + R_1(h), \text{ where } R_1(h) = hf'(a + \theta h).$$

A statement of the purpose of this theorem will help to guide us to its proof. What is desired is a polynomial of specified degree ($n-1$), that approximates the given function as closely as possible in an interval containing a given value of x . Geometrically, we want to approximate the graph of $y = f(x)$ as closely as possible, in the neighborhood of a given point on the graph, by the graph of a polynomial of degree ($n-1$).

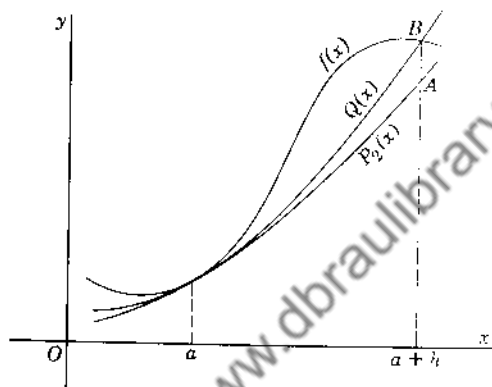


FIG. 1

We shall illustrate a general method of doing this by finding the second-degree polynomial that best approximates the given function in a neighborhood of a point $x = a$, and an expression for the error, or remainder, $R_3(x)$. We shall suppose that $f(x)$ is continuous and has continuous derivatives up to the third order in a neighborhood of $x = a$.

Fixing our attention on the function in the immediate neighborhood of $x = a$, we require: (1) that the second-degree polynomial $P_2(x)$ be equal to $f(x)$ at $x = a$, i.e., $P_2(a) = f(a)$; (2) that $P_2'(a) = f'(a)$, i.e., that the graphs of $P_2(x)$ and $f(x)$ have a common tangent at $x = a$; (3) that $P_2''(a) = f''(a)$, i.e., that the curvatures as well as the slopes of the graphs agree at $x = a$ (Fig. 1). These requirements are sufficient to determine $P_2(x)$, as in Art. 5.

$$P_2(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2. \quad (6.3)$$

We now ask "what is the error in the approximation of $f(x)$ by $P_2(x)$?" That is, for a value distinct from a , say $a+h$, we want an estimate of the difference

$$f(a+h) - P_2(a+h) = f(a+h) - [f(a) + f'(a)h + \frac{1}{2}f''(a)h^2]. \quad (6.4)$$

In Fig. 1, the error is indicated by the line AB .

The mean value theorem gives an expression for the difference between $f(a+h)$ and the constant $f(a)$, in terms of the first derivative of $f(x)$ at a point between a and $a+h$. In the same way, we can now find an expression for the difference between $f(a+h)$ and $P_2(a+h)$ in terms of the third derivative of $f(x)$ at a point between a and $a+h$. For this purpose we make use of a third-degree polynomial $Q(x)$, whose graph agrees with those of $f(x)$ and $P_2(x)$ in height, slope, and curvature at $x=a$, and with the height of $f(x)$ at $x=a+h$. That is,

$$Q(a) = P_2(a) = f(a), \quad Q'(a) = P_2'(a) = f'(a), \\ Q''(a) = P_2''(a) = f''(a), \quad \text{and} \quad Q(a+h) = f(a+h). \quad (6.5)$$

The advantage of this device lies in the fact that $Q(x)$ is known, except for one coefficient.

$$Q(x) = f(a) + f'(a)(x-a) + \frac{1}{2!} f''(a)(x-a)^2 + A(x-a)^3. \quad (6.6)$$

By repeated applications of Rolle's theorem, requiring almost no computation, we can now find A , and hence $Q(x)$, $Q(a+h)$, and $f(a+h) - P_2(a+h)$. Thus,

$f(x) - Q(x) = 0$ at $x=a$ and $x=a+h$; hence, by Rolle's theorem,

$$f'(x) - Q'(x) = 0 \quad \text{for} \quad x = a + \theta_1 h, \quad 0 < \theta_1 < 1.$$

Again, $f'(x) - Q'(x) = 0$ for $x=a$ [by (6.5)] and $x=a+\theta_1 h$; hence, by Rolle's theorem,

$$f''(x) - Q''(x) = 0 \quad \text{for} \quad x = a + \theta_2 h, \quad 0 < \theta_2 < \theta_1 < 1.$$

Finally, $f''(x) - Q''(x) = 0$ for $x=a$ [by (6.5)] and $x=a+\theta_2 h$; hence, by Rolle's theorem,

$$f'''(x) - Q'''(x) = 0 \quad \text{for} \quad x = a + \theta_3 h, \quad 0 < \theta_3 < \theta_2 < \theta_1 < 1.$$

That is,

$$Q'''(a + \theta_3 h) = f'''(a + \theta_3 h). \quad (6.7)$$

Now, from (6.6), $Q'''(x) \equiv 3! A$, and hence, from (6.7),

$$A = \frac{1}{3!} f'''(a + \theta_3 h).$$

Putting this value of A in (6.6),

$$Q(x) = f(a) + f'(a)(x-a) + \frac{1}{2!} f''(a)(x-a)^2 + \frac{1}{3!} f'''(a + \theta_3 h)(x-a)^3.$$

Therefore,

$$Q(a+h) = f(a) + f'(a)h + \frac{1}{2!} f''(a)h^2 + \frac{1}{3!} f'''(a + \theta_3 h)h^3.$$

Since $Q(a+h) = f(a+h)$ and $f(a) + f'(a)h + \frac{1}{2!}f''(a)h^2 = P_2(a+h)$, it follows that the difference

$$f(a+h) - P_2(a+h) = \frac{1}{3!}f'''(a + \theta_3h)h^3.$$

This is the expression sought, for the error (6.4).

If it appears that undue emphasis has been placed on the one point $a+h$, it should be realized that h is a variable and that this point is a typical point in a neighborhood of $x=a$. Conclusions applying to $a+h$ hold for all points in some interval including a . Dropping the subscript from θ_3 , which is no longer needed, we can now write

$$f(a+h) = P_2(a+h) + \frac{1}{3!}f'''(a + \theta h)h^3$$

or

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2!}f''(a)h^2 + \frac{1}{3!}f'''(a + \theta h)h^3. \quad (6.8)$$

Putting h equal to $x-a$,

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + \frac{1}{3!}f'''(x_1)(x-a)^3, \quad (6.9)$$

where x_1 lies between a and x .

Equations 6.8 and 6.9 are statements of Taylor's theorem for three terms and remainder, corresponding to (6.2) and (6.1), respectively. A complete proof for n terms and remainder can be given, similar in all respects to this special case. Since little would be gained by a general proof at this point, we shall omit it. Other more elegant proofs are given in many books, some of which, however, fail to throw light on the meaning of the theorem.

A formal expression of Taylor's theorem, like (6.1) or (6.2), is often called *Taylor's series with remainder*. The relationship between these two forms should be carefully noted. Formula (6.1) expresses $f(x)$ as a polynomial in $x-a$ with remainder containing the factor $f^{(n)}(x_1)$, where x_1 lies between a and x . Here, $x_1 = a + \theta(x-a)$, with $0 < \theta < 1$. Formula 6.2 expresses $f(a+h)$ as a polynomial in powers of h , with remainder containing the factor $f^{(n)}(a + \theta h)$, $a + \theta h$ being a number between a and $a+h$. The formulas are identical if we put $x-a = h$. The letters used are unimportant. Thus, formula 6.2 is often stated with x replacing h , especially if $a=0$, and then it becomes identical with (6.1).

When a function is written as a Taylor's series with remainder it is said to be *expanded* into this form, and the result is called Taylor's *expansion*.

of the function. The conditions for the possibility of such an expansion must not be overlooked. For the validity of (6.1) or (6.2), it is assumed that $f(x)$, $f'(x)$, $f''(x)$, $\dots$, $f^{(n)}(x)$ exist and are continuous at $x = a$. Because of these conditions, one must be careful with such irrational functions as $x^{\frac{1}{2}}$, for example. This function cannot be expanded in powers of x , (i.e., $x - 0$) since none of its derivatives exists for $x = 0$. The same may be said of the transcendental functions $\log x$ and $\cot x$. Similarly, $(1 - x)^{\frac{1}{2}}$ and $\log(1 - x)$ cannot be expanded in powers of $(x - 1)$, nor can $\tan x$ be expanded in powers of $(x - \pi/2)$.

EXAMPLE 1. Find the remainder after 3 terms of the Taylor's series for $\log(1 + x)$ in powers of x , in the form given in (6.1).

SOLUTION. This expansion was found in Art. 2 by a different method and with the remainder in a different form. Following the process embodied in (6.1) with $a = 0$, we write

$$\begin{aligned} f(x) &= \log(1 + x) & f(0) &= 0, \\ f'(x) &= 1/(1 + x) & f'(0) &= 1, \\ f''(x) &= -1/(1 + x)^2 & f''(0) &= -1, \\ f'''(x) &= 2/(1 + x)^3 & f'''(0) &= 2, \\ f^{(4)}(x) &= -3 \cdot 2/(1 + x)^4 & f^{(4)}(0 + \theta x) &= -3!/(1 + 0 + \theta x)^4. \end{aligned}$$

Formally, the first term is 0, and we have, from (6.1),

$$\log(1 + x) = 0 + x - \frac{1}{2}x^2 + R_3(x)$$

where

$$\begin{aligned} R_3(x) &= \frac{1}{3!} f'''(0 + \theta x)x^3 = \frac{2x^3}{3!(1 + 0 + \theta x)^3} \\ &= \frac{x^3}{3(1 + \theta x)^3}. \end{aligned}$$

Here, and in Example 2, θx is identical with x_1 .

The expansion in 3 nonzero terms and remainder is

$$\log(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + R_4(x)$$

where

$$\begin{aligned} R_4(x) &= \frac{1}{4!} f^{(4)}(0 + \theta x)x^4 = -\frac{3!x^4}{4!(1 + 0 + \theta x)^4} \\ &= -\frac{x^4}{4(1 + \theta x)^4}. \end{aligned}$$

EXAMPLE 2. Express $\sqrt{1+x}$ as a polynomial of 3 terms plus a remainder, by Taylor's theorem.

SOLUTION. Differentiating,

$$\begin{aligned} f(x) &= \sqrt{1+x} = (1+x)^{\frac{1}{2}} & f(0) &= 1, \\ f'(x) &= \frac{1}{2}(1+x)^{-\frac{1}{2}} & f'(0) &= \frac{1}{2}, \\ f''(x) &= -\frac{1}{4}(1+x)^{-\frac{3}{2}} & f''(0) &= -\frac{1}{4}, \\ f'''(x) &= \frac{3}{8}(1+x)^{-\frac{5}{2}} & f'''(0+\theta x) &= \frac{3}{8}(1+\theta x)^{-\frac{5}{2}}. \end{aligned}$$

Therefore,

$$\sqrt{1+x} = 1 + \frac{1}{2!}x - \frac{1}{4} \cdot \frac{1}{2!}x^2 + \frac{3}{8} \cdot \frac{1}{3!}x^3(1+\theta x)^{-\frac{5}{2}}.$$

Compare with Example 4, Art. 2.

EXAMPLE 3. Express $(x^2+2)\log x$ in powers of $x-1$, obtaining 4 terms and remainder.

SOLUTION. We use (6.1) with $a=1$. Differentiating,

$$\begin{aligned} f(x) &= (x^2+2)\log x & f(1) &= 0, \\ f'(x) &= x + 2/x + 2x\log x & f'(1) &= 3, \\ f''(x) &= 3 - 2/x^2 + 2\log x & f''(1) &= 1, \\ f'''(x) &= 4/x^3 + 2/x & f'''(1) &= 6, \\ f^{(4)}(x) &= -12/x^4 - 2/x^2 & f^{(4)}(1) &= -14, \\ f^{(5)}(x) &= 48/x^5 + 4/x^3 & f^{(5)}(x_1) &= 48/x_1^5 + 4/x_1^3. \end{aligned}$$

Hence,

$$\begin{aligned} (x^2+2)\log x &= 3(x-1) + \frac{1}{2!}(x-1)^2 + \frac{1}{3!} \cdot 6(x-1)^3 \\ &\quad + \frac{1}{4!}(-14)(x-1)^4 + \frac{1}{5!}(x-1)^5(48/x_1^5 + 4/x_1^3), \end{aligned}$$

in which x_1 lies between 1 and x , i.e., $x_1 = 1 + \theta(x-1)$.

EXERCISES

Obtain the expansions in Exercises 1-10.

$$1. e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 e^{\theta x}.$$

$$2. \cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \frac{1}{8!}x^8 \cos(\theta x).$$

$$3. \cosh x = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \frac{1}{6!}x^6 + \frac{1}{8!}x^8 \cosh(\theta x).$$

$$4. \sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \frac{1}{9!}x^9 \cos(\theta x).$$

$$5. \sinh x = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \frac{1}{7!}x^7 + \frac{1}{9!}x^9 \cosh(\theta x).$$

$$6. \sin x + \cos x = 1 + x - \frac{1}{2!}x^2 - \frac{1}{3!}x^3 + \frac{1}{4!}x^4 [\sin(\theta x) + \cos(\theta x)].$$

$$7. (1+x)^{\frac{1}{2}} = 1 + \frac{1}{2}x - \frac{2}{25}x^2 + \frac{6}{125}x^3 - \frac{42}{1250}x^4 (1+\theta x)^{-\frac{1}{2}}.$$

$$8. 1/(1-x)^2 = 1 + 2x + 3x^2 + 4x^3 + 5x^4/(1-\theta x)^6.$$

$$9. (16+x)^{\frac{1}{2}} = 2 \left(1 + \frac{1}{16}x \right)^{\frac{1}{2}} \\ = 2 \left[1 + \frac{1}{64}x - \frac{3}{2(16)^3}x^2 + \frac{7}{8(16)^4}x^3 \left(1 + \frac{1}{16}\theta x \right)^{-\frac{1}{2}} \right].$$

$$10. \tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{1}{315}x^7 [17 + 248 \tan^2(\theta x) + 756 \tan^4(\theta x) \\ + 840 \tan^6(\theta x) + 315 \tan^8(\theta x)].$$

$$11. \text{Show that if } x < 1, \sin x = x - x^3/3! \text{ with an error of less than } .01.$$

12. Why is a Taylor's expansion in powers of x impossible for the function $\log(e^x - 1)$?

13. Expand $\log \cos x$ in powers of x into two nonzero terms and remainder.

14. Obtain $\log(1+e^x) = \log 2 + \frac{1}{2}x + \frac{1}{4}x^2 - \frac{1}{192}x^4 + A$, and find an expression for A .

$$15. \text{Obtain } e^{2x} = e^2 [1 + 2(x-1) + 2(x-1)^2 + \frac{4}{3}(x-1)^3 + \frac{2}{3}(x-1)^4 e^{2\theta x}].$$

$$16. \text{Obtain } \sin(3x - \pi/2) = -1 + \frac{9}{2!}x^2 - \frac{81}{4!}x^4 + \frac{729}{6!}x^6 \cos(3\theta x).$$

$$17. \text{Obtain } \sin(3x - \pi/2) = -3(x - \pi/2) + \frac{27}{3!}(x - \pi/2)^3 - \frac{243}{5!}(x - \pi/2)^5 \\ - \frac{2187}{7!}(x - \pi/2)^7 \sin 3x_1, \text{ where } x_1 \text{ lies between } \pi/2 \text{ and } x.$$

$$18. \text{Expand } e^{-x} \text{ to 4 terms and remainder in powers of } (x+1). \text{ [Hint: } a = -1.]$$

$$19. \text{Obtain } 1/x^3 = \frac{1}{8} - \frac{3}{16}(x-2) + \frac{3}{16}(x-2)^2 - \frac{5}{82}(x-2)^3 + \frac{15}{15}(x-2)^4/x_1^7, \\ \text{where } x_1 \text{ lies between 2 and } x.$$

$$20. \text{Obtain } \tan^{-1} x = \pi/4 + \frac{1}{2}(x-1) - \frac{1}{4}(x-1)^2 + \frac{1}{12}(x-1)^3 \\ + \frac{1}{8}(x-1)^4(1 - 10x_1^2 + 5x_1^4)/(1 + x_1^2)^6, \text{ where } x_1 \text{ lies between 1 and } x.$$

7. Numerical Approximations. In Chapter 12 we saw how the mean value theorem could be used to make numerical approximations and give estimates of maximum errors. Since Taylor's theorem is really an

extension of the mean value theorem, it can be used for the same purpose. Approximations as close as desired will be possible by Taylor's theorem, whenever $R_n \rightarrow 0$ as $n \rightarrow \infty$.

EXAMPLE 1. Approximate $e^{.2}$ with an error less than .0001, by Taylor's theorem.

SOLUTION. We let $f(x) = e^x$ and write the Taylor's expansion: $e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{(n-1)!}x^{n-1} + R_n(x)$, in which $R_n(x) = \frac{1}{n!}x^n e^{\theta x}$. We now put $x = .2$, and $R_n(x)$ becomes $R_n = \frac{1}{n!}(.2)^n e^{.2\theta}$. Now, $e^{.2\theta} < e^{\frac{1}{2}} < \frac{5}{4}$, as we see by raising $\frac{5}{4}$ to the fifth power: $(\frac{5}{4})^5 = \frac{3125}{1024} > 3 > e$. Hence, $R_n < \frac{1}{n!}(.2)^n \frac{5}{4}$. We find by experiment that if $n = 4$, $R_n < .000083$. Therefore,

$$\begin{aligned} e^{.2} &= 1 + .2 + \frac{1}{2!}(.2)^2 + \frac{1}{3!}(.2)^3, \text{ with error } < .000083 \\ &= 1.2213, \text{ with error } < .000086. \end{aligned}$$

(The last digit includes the error in rounding off the quotient in $\frac{1}{3!}(.2)^3$.)

EXAMPLE 2. Approximate $\sqrt[4]{18}$ to 3 decimal places by Taylor's theorem.

SOLUTION. We write $\sqrt[4]{18} = (16 + 2)^{\frac{1}{4}} = [16(1 + \frac{1}{8})]^{\frac{1}{4}} = 2(1 + \frac{1}{8})^{\frac{1}{4}}$. We now identify $1 + \frac{1}{8}$ with $a + h$, i.e., $a = 1$, $h = \frac{1}{8}$, and put $f(a + h) = (a + h)^{\frac{1}{4}}$, $f(a) = a^{\frac{1}{4}}$. Now, for the function $f(x) = x^{\frac{1}{4}}$, all derivatives exist and are continuous except at $x = 0$, and hence are continuous at $x = a = 1$. Applying Taylor's theorem with 3 terms and remainder, at the start,

$$(a + h)^{\frac{1}{4}} = a^{\frac{1}{4}} + \frac{1}{4}a^{-\frac{3}{4}}h - \frac{3}{32}a^{-\frac{5}{4}}h^2 + \frac{7}{128}h^3(a + \theta h)^{-\frac{7}{4}}.$$

Putting $a = 1$, $h = \frac{1}{8}$, we have

$$(1 + \frac{1}{8})^{\frac{1}{4}} = 1 + \frac{1}{32} - \frac{3}{2048} + \frac{7}{65536} \cdot \frac{1}{(1 + \frac{1}{8}\theta)^{\frac{7}{4}}}.$$

Giving θ its minimum value of 0, we see that the remainder after 3 terms is less than $7/65,000$, which is in the neighborhood of .0001 (actually .000108). It follows that $(1 + \frac{1}{8})^{\frac{1}{4}} = 1 + \frac{1}{32} - \frac{3}{2048}$, with an error less than .0002, and $2(1 + \frac{1}{8})^{\frac{1}{4}} = 2(1 + \frac{1}{32} - \frac{3}{2048})$ with an error less than .0004. Hence, $\sqrt[4]{18} = 2.060$, correct to 3 decimal places.

EXERCISES

Use Taylor's theorem to make the following approximations.

1. $\sin(\frac{1}{2} \text{ rad})$, with an error less than .000002. [Hint: $|R_n| < |h^n|/n!$ since $f^{(n)}(a + \theta h)$ is either the sine or the cosine of some quantity, and hence ≤ 1 .]
2. $\sqrt[3]{68}$, with an error less than .00001.
3. $1/\sqrt[3]{36}$, with an error less than .0005, i.e., to 3 decimal places.
4. $e^{1.2}$, with an error less than .0005, and hence to 3 decimal places. [Hint: Use the expansion: $e^{a+h} = e^a \left[1 + h + \frac{1}{2!} h^2 + \cdots + \frac{1}{(n-1)!} h^{n-1} + \frac{1}{n!} h^n e^{a+\theta h} \right]$
5. $\cosh 1$, to 4 decimal places.

8. L'Hospital's Rule. It is the nature of a derivative that it is a limit of a ratio, both numerator and denominator of which tend to 0 together. All differentiation consists of methods of obtaining such limits. Such ratios occur in other connections, as well, and it is desirable to have convenient methods of obtaining their limits. One such method can be devised with the help of Taylor's theorem.

Let $f(x)$ and $g(x)$ be two functions which are continuous at $x = a$ and possess continuous derivatives at that point, and let $f(a) = g(a) = 0$. Then the function $\phi(x) = f(x)/g(x)$ is discontinuous at $x = a$. Our problem is to find the limit of $\phi(x)$ as $x \rightarrow a$, if it exists. For this purpose we can write, by the mean value theorem,

$$\phi(a+h) = \frac{f(a+h)}{g(a+h)} = \frac{f(a) + hf'(a+\theta h)}{g(a) + hg'(a+\theta' h)},$$

where both θ and θ' lie between 0 and 1. Since $f(a) = g(a) = 0$, we have

$$\phi(a+h) = \frac{hf'(a+\theta h)}{hg'(a+\theta' h)}$$

and

$$\lim_{h \rightarrow 0} \phi(a+h) = \lim_{h \rightarrow 0} \frac{f'(a+\theta h)}{g'(a+\theta' h)}.$$

Applying the theorem on the limit of a quotient, if $g'(a) \neq 0$ we get

$$\lim_{h \rightarrow 0} \phi(a+h) = \frac{\lim_{h \rightarrow 0} f'(a+\theta h)}{\lim_{h \rightarrow 0} g'(a+\theta' h)} = \frac{f'(a)}{g'(a)}. \quad (8.1)$$

If $g'(a) = 0$ and $f'(a) \neq 0$, then $\phi(a+h) \rightarrow \infty$ as $h \rightarrow 0$. If $g'(a) = f'(a) = 0$, Eq. 8.1 does not yield the desired limit. In this event, if $f(x)$ and $g(x)$ have continuous second derivatives at $x = a$, we write

$$\phi(a+h) = \frac{f(a+h)}{g(a+h)} = \frac{f(a) + hf'(a) + \frac{h^2}{2} f''(a+\theta_1 h)}{g(a) + hg'(a) + \frac{h^2}{2} g''(a+\theta_1' h)}.$$

Since $f(a) = g(a) = f'(a) = g'(a) = 0$,

$$\phi(a + h) = \frac{h^2 f''(a + \theta_1 h)}{h^2 g''(a + \theta_1' h)},$$

and

$$\lim_{h \rightarrow 0} \phi(a + h) = \frac{f''(a)}{g''(a)}, \quad \text{provided that } g''(a) \neq 0.$$

This method can be extended without difficulty, leading to *L'Hospital's rule*, which can be stated as follows.

Let $f(x)$ and $g(x)$ possess continuous derivatives of all orders up to the n th; let $f(a) = g(a) = 0$, and let the lowest order derivative of either function which fails to vanish at $x = a$ be the n th. There are then three possibilities regarding the limit of the ratio $f(x)/g(x)$ as $x \rightarrow a$, as follows.

1. If $f^{(n)}(a) \neq 0$ and $g^{(n)}(a) \neq 0$, then $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{f^{(n)}(a)}{g^{(n)}(a)}$;
2. If $f^{(n)}(a) = 0$ and $g^{(n)}(a) \neq 0$, then $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$;
3. If $f^{(n)}(a) \neq 0$ and $g^{(n)}(a) = 0$, then $\frac{f(x)}{g(x)} \rightarrow \infty$ as $x \rightarrow a$.

In practice, we may write

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow a} \frac{f''(x)}{g''(x)} = \cdots,$$

discontinuing the process as soon as the limit of either the numerator or the denominator is different from 0.

Since $f(a)$ and $g(a) = 0$, the evaluation of $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$ is sometimes confusingly called the "evaluation of the indeterminate form $0/0$." This terminology is obviously misleading, since one cannot expect to evaluate that which is really indeterminate, and since the meaningless symbol $0/0$ is certainly not given a meaning or a value by the process used and the result obtained.

L'Hospital's rule can also be used to evaluate $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$ if $f(x)$ and $g(x)$ both tend to infinity as x tends to a , in exactly the same way as when $f(x)$ and $g(x)$ both tend to 0. We shall not prove this fact. This application is illustrated by Example 2.

Although L'Hospital's rule frequently provides a convenient way of dealing with limits of this kind, it should not be relied upon exclusively.

Often, it is preferable to make direct use of known limits. For convenience, some of these are listed below.

$$\lim_{x \rightarrow \infty} \frac{a_0 x^n + a_1 x^{n-1} + \cdots + a_n}{b_0 x^n + b_1 x^{n-1} + \cdots + b_n} = \frac{a_0}{b_0} \quad (8.2)$$

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1. \quad (8.3)$$

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x} = 0. \quad (8.4)$$

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}. \quad (8.5)$$

$$\lim_{x \rightarrow 0} \frac{\sinh x}{x} = 1. \quad (8.6)$$

$$\lim_{x \rightarrow 0} \frac{\cosh x - 1}{x} = 0. \quad (8.7)$$

$$\lim_{x \rightarrow 0} \frac{\cosh x - 1}{x^2} = \frac{1}{2}. \quad (8.8)$$

$$\lim_{x \rightarrow \infty} \frac{\log x}{x^a} = 0, \quad \text{if } a > 0. \quad (8.9)$$

$$\lim_{x \rightarrow \infty} \frac{x^a}{b^x} = 0, \quad \text{if } b > 1. \quad (8.10)$$

$$\lim_{x \rightarrow 0} x^a \log x = 0, \quad \text{if } a > 0. \quad (8.11)$$

$$\lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \log a, \quad \text{if } a > 0. \quad (\text{This is } f'(0), \text{ where } f(x) = a^x.) \quad (8.12)$$

$$\lim_{x \rightarrow 0} \frac{\log(1+x)}{x} = 1. \quad (\text{This is } f'(1), \text{ where } f(x) = \log x.) \quad (8.13)$$

The following examples show the adaptations of L'Hospital's rule to functions of several different forms.

EXAMPLE 1. Find $\lim_{x \rightarrow 0} \frac{\sin x - x}{x - \tan x}$.

SOLUTION. Since both numerator and denominator are 0 when $x = 0$, we can apply L'Hospital's rule.

$$\lim_{x \rightarrow 0} \frac{\sin x - x}{x - \tan x} = \lim_{x \rightarrow 0} \frac{\cos x - 1}{1 - \sec^2 x}$$

Now,

$$\frac{\cos x - 1}{1 - \sec^2 x} = \frac{(\cos x - 1) \cos^2 x}{\cos^2 x - 1} = \frac{\cos^3 x}{\cos x - 1},$$

and the latter fraction $\rightarrow \frac{1}{2}$ as $x \rightarrow 0$.

EXAMPLE 2. Find $\lim_{x \rightarrow 0} \frac{\log x}{\operatorname{ctn} x}$.

SOLUTION. Here $\log x \rightarrow -\infty$ and $\operatorname{ctn} x \rightarrow \infty$ as $x \rightarrow 0$. Applying the rule, we get

$$\lim_{x \rightarrow 0} \frac{\log x}{\operatorname{ctn} x} = \lim_{x \rightarrow 0} \frac{1/x}{-\csc^2 x}.$$

This fraction can be written $-\frac{\sin^2 x}{x}$, and the limit is clearly 0, since $\lim_{x \rightarrow 0} \frac{\sin x}{x} \cdot \sin x = 1 \cdot 0$.

This limit could have been found by direct use of (8.3) and (8.11) as follows.

$$\frac{\log x}{\operatorname{ctn} x} = \frac{\sin x}{\cos x} \log x = \frac{1}{\cos x} \cdot \frac{\sin x}{x} \cdot x \log x.$$

Hence,

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\log x}{\operatorname{ctn} x} &= \lim_{x \rightarrow 0} \left[\frac{1}{\cos x} \cdot \frac{\sin x}{x} \cdot (x \log x) \right] \\ &= 1 \cdot 1 \cdot 0 = 0. \end{aligned}$$

EXAMPLE 3. Find $\lim_{x \rightarrow 0} \left(\frac{2}{x^2} - \frac{1}{1 - \cos x} \right)$.

SOLUTION. Although each fraction tends to ∞ as x tends to 0, it is possible that their difference may tend to a finite limit. By combining them into a single fraction we obtain a function to which L'Hospital's rule can be applied. Thus,

$$\begin{aligned} \lim_{x \rightarrow 0} \left(\frac{2}{x^2} - \frac{1}{1 - \cos x} \right) &= \lim_{x \rightarrow 0} \frac{2 - 2 \cos x - x^2}{x^2(1 - \cos x)} \\ &= \lim_{x \rightarrow 0} \frac{2 \sin x - 2x}{2x(1 - \cos x) + x^2 \sin x} \\ &= \lim_{x \rightarrow 0} \frac{2(\cos x - 1)}{2(1 - \cos x) + 4x \sin x + x^2 \cos x} \end{aligned}$$

Dividing both numerator and denominator of this fraction by x^2 , we have

$$\frac{2(\cos x - 1)}{x^2} \bigg/ \left[\frac{2(1 - \cos x)}{x^2} + \frac{4 \sin x}{x} + \cos x \right].$$

Now, using (8.5) and (8.3), we find the limit to be

$$-1/(1 + 4 + 1) = -1/6.$$

EXAMPLE 4. Find $\lim_{x \rightarrow \pi} \frac{\sin x + x - \pi}{1 + \cos x}$.

SOLUTION. By L'Hospital's rule, this limit is equal to

$$\lim_{x \rightarrow \pi} \frac{\cos x + 1}{-\sin x} = \lim_{x \rightarrow \pi} \frac{-\sin x}{-\cos x} = 0.$$

EXAMPLE 5. Find $\lim_{x \rightarrow 0} (1 + \sin x)^{1/x}$.

SOLUTION. As $x \rightarrow 0$, $(1 + \sin x) \rightarrow 1$, while $1/x \rightarrow \infty$. This form can be brought under L'Hospital's rule by a simple device. We set $y = (1 + \sin x)^{1/x}$ and write $\log y = \frac{\log(1 + \sin x)}{x}$. If we can find the limit of $\log y$, we can find the limit of y itself. Now

$$\lim_{x \rightarrow 0} \frac{\log(1 + \sin x)}{x} = \lim_{x \rightarrow 0} \frac{(\cos x)/(1 + \sin x)}{1} = 1.$$

Hence,

$$\lim_{x \rightarrow 0} \log y = 1, \text{ and } \lim_{x \rightarrow 0} y = e.$$

EXERCISES

Find the limits in Exercises 1-20 by means of L'Hospital's rule.

$$1. \lim_{x \rightarrow 0} \frac{\sin x - x}{\cos x - 1}$$

$$2. \lim_{x \rightarrow 0} \frac{\sinh x - x}{x - \sin x}$$

$$3. \lim_{x \rightarrow 1} \frac{\log x}{x - 1}$$

$$4. \lim_{x \rightarrow \pi} \frac{\log \sin(x/2)}{x - \pi}$$

$$5. \lim_{x \rightarrow \pi} \frac{\log \sin(x/2)}{(x - \pi)^2}$$

$$6. \lim_{x \rightarrow 0} \frac{x - \sin x}{x^3}$$

$$7. \lim_{x \rightarrow 0} \frac{x - \sin x}{\tan x - x}$$

$$8. \lim_{x \rightarrow 0} \frac{x - \sin^{-1} x}{\sin^3 x}$$

$$9. \lim_{x \rightarrow \infty} \frac{x + \log x}{x \log x}$$

$$10. \lim_{x \rightarrow \pi/2} \frac{\log(x - \pi/2)}{\log \sec x}$$

11. $\lim_{x \rightarrow \pi/2} \frac{\log(x - \pi/2)}{\tan x}$

12. $\lim_{x \rightarrow 0} x^x$

13. $\lim_{x \rightarrow 0} x^{\sin x}$

14. $\lim_{x \rightarrow \pi} e^x \cot x$

15. $\lim_{x \rightarrow 0} (x \log \sin x)$

16. $\lim_{x \rightarrow 0} \left(\frac{1}{x} - \frac{1}{\sin x} \right)$

17. $\lim_{x \rightarrow \pi/2} [x \sec x - (\pi/2) \tan x]$

18. $\lim_{x \rightarrow 1} x^{1/(x-1)}$

19. $\lim_{x \rightarrow 0} (1 + x)^{1/x} \csc x$

20. $\lim_{x \rightarrow \pi/2} (\sin x)^{\tan x}$

Find the limits in Exercises 21–35 without using L'Hospital's rule.

21. $\lim_{x \rightarrow 3} \frac{x^2 - 9}{x^2 - x - 6}$

22. $\lim_{x \rightarrow \infty} \frac{5x^2 + 8}{2x^2 + x + 1}$

23. $\lim_{x \rightarrow 0} \frac{\tan 5x}{\tan x}$

24. $\lim_{x \rightarrow 0} \frac{\sinh x}{\sin x}$

25. $\lim_{x \rightarrow 0} \frac{1 - \sqrt{1 - x^2}}{x^2}$

26. $\lim_{x \rightarrow \infty} \frac{\log x - \sqrt{x}}{x}$

27. $\lim_{x \rightarrow 0} \frac{e^x + e^{-x} - 2 \cos x}{x \sin x}$

28. $\lim_{x \rightarrow 0} (1 + x)^{1/x}$

29. $\lim_{x \rightarrow 1} x^{1/(1-x)}$

30. $\lim_{x \rightarrow 0} \frac{e^{-1/x}}{x^n}$

31. $\lim_{x \rightarrow \infty} e^{-x} \log x$

32. $\lim_{x \rightarrow \infty} x^n e^{-x} \log x$

33. $\lim_{x \rightarrow 1} \frac{\sin \pi x}{1 - x}$

34. $\lim_{x \rightarrow \pi/4} \frac{\tan x - 1}{x - \pi/4}$

35. $\lim_{x \rightarrow 0} \frac{\csc x - \cot x}{x}$

Find the limits in Exercises 36–45 by any method.

36. $\lim_{x \rightarrow 0} \frac{(1+x)^{1/n} - 1}{x}$

37. $\lim_{x \rightarrow \infty} \sqrt[5]{\frac{32x^5 + 1}{243x^5 + 1}}$

38. $\lim_{x \rightarrow a} \frac{x - a + \sqrt{x^{11} - a^{11}}}{\sqrt{x^7 - a^7}}$

39. $\lim_{x \rightarrow \infty} \frac{x - \sin x}{x}$

40. $\lim_{x \rightarrow 0} \left[\frac{1}{x} - \frac{1}{x^2} \log(1+x) \right]$

41. $\lim_{x \rightarrow \pi/2} \frac{\tan x}{\log(x - \pi/2)}$

42. $\lim_{x \rightarrow 0} \frac{\sinh x + \sin x - 2x}{x^5}$

43. $\lim_{x \rightarrow 0} (\cosh x)^{1/x}$

44. $\lim_{x \rightarrow 0} \frac{\log(x + \sqrt{1+x^2})}{x}$

45. $\lim_{x \rightarrow 1} (\sqrt{1-x} \cdot \tanh^{-1} x)$

9. Summary. Our purpose in this chapter was to discover polynomial approximations to more or less arbitrary functions. This led to the development of Taylor's theorem, the importance of which greatly transcends our original purpose. In fact, this theorem has been called by some the most important single theorem in calculus. In addition to the uses of this theorem that were made in the present chapter, it will occur again in Chapter 22 as the foundation on which will rest much of the theory of power series.

Downloaded from www.dbraulibrary.org.in

Infinite Series of Constants

1. Introduction. The tangent problem and the area problem are the characteristic problems of calculus. Each is a particular application of the concept of the limit of a sequence. We now consider a third application of this concept which is most useful in the further development of calculus, and which will prove to be essentially simpler to understand than the two limits with which we have been almost exclusively occupied up to this time. This application is to the sum of an infinite series of terms.

Our attention is drawn to a sum of this kind by the famous paradox of the race-course, attributed to Zeno, the Stoic philosopher. This paradox can be stated as follows. A runner can never reach the end of a race-course, since, before he can traverse the whole course he must run one-half of the course; he must then run one-half of the remaining half; then one-half of the part still remaining; and so on. Since there are infinitely many parts of the racecourse to be run, he cannot reach the end in a finite time.

In order to resolve this paradox, we set it up in numerical form. Suppose, for example, the runner runs on a 1-mile course at a constant speed of 12 mph, whereupon he ought to run the length of the course in 5 min. The infinitely many parts of the course mentioned above have the lengths $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, $\frac{1}{32}$, etc., and their sum is $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \dots$, with the dots indicating all the infinitely many remaining terms. Clearly, the sum of any finite number of these terms is less than 1, since there is always $\frac{1}{2}$ of some remaining part of the 1-mile course yet to be added before arriving at 1. But, since the part remaining is less and less, and approaches 0 as more and more terms are added, one might be led to conclude that the sum of *all* the terms is exactly 1. In the same way, the time required to run $\frac{1}{2}$ mile is $\frac{1}{2} \cdot 5$ min; for $\frac{1}{4}$ mile it is $\frac{1}{4} \cdot 5$ min; for $\frac{1}{8}$ mile it is $\frac{1}{8} \cdot 5$ min; etc. Hence, the sum of the intervals of time corresponding to the several parts of the distance is

$$\frac{1}{2} \cdot 5 + \frac{1}{4} \cdot 5 + \frac{1}{8} \cdot 5 + \frac{1}{16} \cdot 5 + \frac{1}{32} \cdot 5 + \dots$$

Intuitively, we conclude that all these terms must add up to 5, so that the end of the racecourse is reached in 5 minutes, as expected.

We are thus led to believe that some infinite series can have perfectly definite finite sums. Of course, some do not. For example, the series $1 - 2 + 3 - 4 + 5 - \dots$, cannot have a finite sum. It is natural to ask how one can distinguish between infinite series that have finite sums and those that do not. As a first step, we must define the sum of an infinite series. It should be clear that a new definition is required, since our definition of the sum, on which addition in algebra and arithmetic is based, has meaning only for the sum of a finite number of quantities.

2. The Sum of an Infinite Series. A definition of the sum of an infinite series, satisfying all intuitive requirements, can be based upon the concept of the limit of a sequence. Let us illustrate the procedure by applying it to the series of distances in Art. 1: $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \dots$. We form the infinite sequence of partial sums:

$$S_1 = \frac{1}{2}, \quad S_2 = \frac{1}{2} + \frac{1}{4}, \quad S_3 = \frac{1}{2} + \frac{1}{4} + \frac{1}{8}, \quad \dots$$

$$S_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots + \frac{1}{2^n},$$

$\dots$

Now, since S_n is the sum of a geometric progression with first term $= \frac{1}{2}$, common ratio $= \frac{1}{2}$, and with n terms, its value is

$$S_n = \frac{\frac{1}{2} - \frac{1}{2} \cdot (\frac{1}{2})^n}{1 - \frac{1}{2}} = 1 - \frac{1}{2^n}$$

and $\lim_{n \rightarrow \infty} S_n = 1$. Since $\lim_{n \rightarrow \infty} S_n$ is the limit approached by the sum of more and more terms of the series, it is natural to define the sum of the whole series by this limit. This answer agrees with the intuitive answer given in Art. 1.

A general expression for an infinite series is

$$a_1 + a_2 + a_3 + \dots + a_n + \dots$$

The partial sums of this general series are

$$S_1 = a_1, \quad S_2 = a_1 + a_2, \quad S_3 = a_1 + a_2 + a_3, \quad \dots$$

$$S_n = a_1 + a_2 + a_3 + \dots + a_n,$$

$\dots$

We now define the sum of the infinite series to be

$$S = \lim_{n \rightarrow \infty} S_n, \text{ if this limit exists.}$$

If the sequence $S_1, S_2, \dots, S_n, \dots$ has a limit, the infinite series is said to *converge*. Otherwise, it *diverges*.

The definition of the sum of an infinite series is also expressed formally by the equation

$$a_1 + a_2 + \cdots \text{ to infinitely many terms} = \lim_{n \rightarrow \infty} (a_1 + a_2 + \cdots + a_n),$$

if this limit exists. Employing the summation symbol Σ , we denote the sum of an infinite series

$$a_1 + a_2 + \cdots \text{ to infinitely many terms, by } \sum_{i=1}^{\infty} a_i.$$

The definition of the sum can therefore be expressed

$$S = \sum_{i=1}^{\infty} a_i = \lim_{n \rightarrow \infty} \sum_{i=1}^n a_i, \text{ if this limit exists.}$$

The terms at the beginning of a series have nothing to do with the question of its convergence or divergence. This follows from the fact that the early terms in the sequence of partial sums have nothing to do with the convergence of the sequence; it is what happens to S_n as n increases beyond all bounds that is decisive. Of course, the addition or subtraction of a finite number of terms at the beginning can change the sum to which a series converges, but the fact of convergence or divergence is unchanged thereby.

3. Examples of Infinite Series. The series

$$\sum_{i=1}^{\infty} 1/2^i = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots,$$

which was discussed in Art. 2, is an example of an infinite geometric series. Its sum, obtained in accordance with the definition, is, as we learned,

$$\sum_{i=1}^{\infty} 1/2^i = \lim_{n \rightarrow \infty} \sum_{i=1}^n 1/2^i = \lim_{n \rightarrow \infty} (1 - 1/2^n) = 1.$$

Let us consider the general geometric series

$$\sum_{i=1}^{\infty} ar^{i-1} = a + ar + ar^2 + \cdots.$$

The partial sums are

$$S_1 = a, \quad S_2 = a + ar, \quad S_3 = a + ar + ar^2, \quad S_n = \sum_{i=1}^n ar^{i-1}.$$

From the formula for the sum of a finite geometric progression, it follows that

$$S_n = \frac{a - ar^n}{1 - r} = \frac{a}{1 - r} - \frac{ar^n}{1 - r}. \quad (3.1)$$

From Art. 6, Chapter 2, we have

$$\lim_{n \rightarrow \infty} ar^n = 0 \quad \text{if } r < 1,$$

$$= a \quad \text{if } r = 1,$$

and

$$ar^n \rightarrow \infty \quad \text{if } r > 1.$$

Hence,

$$\lim_{n \rightarrow \infty} S_n = \frac{a}{1-r}, \quad \text{if } r < 1,$$

whereas S_n has no limit if $r > 1$. If $r = 1$, formula 3.1 ceases to apply, since the denominator is 0, but we see that

$$S_n = a + a + a + \cdots + a = na,$$

and hence $S_n \rightarrow \infty$. Thus, the infinite geometric series converges to $a/(1-r)$ if the common ratio is less than 1, and diverges otherwise.

It is, of course, obvious that every arithmetic series,

$$a + (a+d) + (a+2d) + (a+3d) + \cdots$$

diverges, since the numerical values of the terms eventually surpass all bounds. More generally, it is a necessary condition for the convergence of a series that, eventually, the terms of the series become very small and approach 0 as a limit as n tends to ∞ . For, the fact that the sequence of partial sums converges to a limit S requires that both S_n and S_{n-1} tend to S as $n \rightarrow \infty$. Since $a_n = S_n - S_{n-1}$, it follows that

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (S_n - S_{n-1}) = \lim_{n \rightarrow \infty} S_n - \lim_{n \rightarrow \infty} S_{n-1} = S - S = 0.$$

That this condition is not sufficient to make a series converge is shown by the following example, in which the terms eventually tend to 0, although the series diverges.

We consider the harmonic series

$$\sum_{i=1}^{\infty} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \cdots + \frac{1}{n} + \cdots \quad (3.2)$$

We compare the partial sums of this series with those of the series

$$\frac{1}{2} + \frac{1}{2} + [\frac{1}{4} + \frac{1}{4}] + [\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}]$$

$$+ [\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}] + \cdots \quad (3.3)$$

which is obtained from the harmonic series by replacing 1 by $\frac{1}{2}$; $\frac{1}{3}$ by $\frac{1}{4}$; $\frac{1}{5}$, $\frac{1}{6}$, $\frac{1}{7}$, each by $\frac{1}{8}$; $\frac{1}{9}$ to $\frac{1}{15}$ each by $\frac{1}{16}$; etc. Series 3.3 diverges since, by grouping the terms as indicated by the brackets, it is equal to $\frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots$. If we let S_n denote the partial sums of the

harmonic series 3.2 and S_n' denote the partial sums of series 3.3, we have $S_n > S_n'$, for every value of n . But the terms of S_n' increase beyond all bounds:

$$S_1' = \frac{1}{2}, \quad S_2' = 2 \cdot \frac{1}{2}, \quad S_3' = 3 \cdot \frac{1}{2}, \quad S_4' = 4 \cdot \frac{1}{2}, \text{ etc.}$$

Hence, the terms of the sequence S_n also increase beyond all bounds, and series 3.2 diverges.

EXERCISES

1. Write the first 5 terms of each of the series.

$$(a) \sum_{i=1}^{\infty} \frac{1}{i^2}$$

$$(b) \sum_{i=1}^{\infty} \frac{2^{i-1}}{\sqrt{i}}$$

$$(c) \sum_{i=1}^{\infty} \frac{1}{i!}$$

$$(d) \sum_{i=1}^{\infty} \frac{i}{2^i}$$

$$(e) \sum_{i=1}^{\infty} \frac{i^{i-1}}{(2i)!}$$

$$(f) \sum_{i=1}^{\infty} \frac{(i-1) \cdot (i-2) \cdots (i-1)!}{(i-1)!}$$

2. By inspection, find a general law of formation of each of the following series, i.e. a formula for a_i .

$$(a) \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$$

$$(b) 1 - \frac{1}{3} + \frac{1}{9} - \frac{1}{27} + \cdots$$

$$(c) 0 + \frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \cdots$$

$$(d) \frac{1}{1 \cdot 2} + \frac{3}{2 \cdot 3} + \frac{5}{3 \cdot 4} + \frac{7}{4 \cdot 5} + \cdots$$

$$(e) \frac{1}{1} - \frac{2}{1 \cdot 2} + \frac{4}{1 \cdot 2 \cdot 3} - \frac{8}{1 \cdot 2 \cdot 3 \cdot 4} + \cdots$$

$$(f) \frac{a}{2} + \frac{a\sqrt{a}}{2 \cdot 4} + \frac{a^2}{2 \cdot 4 \cdot 6} + \frac{a^2\sqrt{a}}{2 \cdot 4 \cdot 6 \cdot 8} + \cdots$$

3. Verify the following infinite sums.

$$(a) 1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \cdots = \frac{3}{2}$$

$$(b) 1 + \frac{9}{10} + \frac{81}{100} + \frac{729}{1000} + \cdots = 10$$

$$(c) .14 + .0014 + .000014 + .00000014 + \cdots = \frac{1}{70}$$

$$(d) -2 + 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \cdots = -\frac{4}{3}$$

4. Show that the harmonic series $\sum_{i=1}^{\infty} \frac{1}{2i+1}$ diverges.

5. Show that $\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n(n+1)} + \cdots$ converges to the limit 1. [Hint: Express $1/[i \cdot (i+1)]$ in partial fractions.]

6. Show that $\frac{3}{1 \cdot 2} - \frac{5}{2 \cdot 3} + \frac{7}{3 \cdot 4} - \cdots + (-1)^{n-1} \frac{2n-1}{n(n+1)} + \cdots$ converges to the limit 1. [Hint: Break up into partial fractions, as in Exercise 5.]

7. Show that $\frac{1}{1 \cdot 2 \cdot 3} + \frac{1}{2 \cdot 3 \cdot 4} + \frac{1}{3 \cdot 4 \cdot 5} + \cdots + \frac{1}{n(n+1)(n+2)} + \cdots$ converges to the limit $\frac{1}{2}$. [Hint: Break up into partial fractions, as in Exercise 5.]

4. Two Theorems on Infinite Sequences. The definition of the sum of an infinite series shows that the theory of infinite series, like differentiation and integration, rests fundamentally on the theory of sequences. In order to proceed, unhampered, in our study of series, we need two basic theorems on sequences not formerly stated. The first of these is called the Cauchy* test, or Cauchy intrinsic principle of convergence. The adjective "intrinsic" refers to the fact that, by this test, the convergence or divergence of a sequence is determined from knowledge contained within the sequence, without prior knowledge of the limit.

THEOREM (CAUCHY TEST). A sequence $S_1, S_2, \dots$ converges to a limit if, and only if, for each positive ε there exists a positive integer N , such that $|S_n - S_m| < \varepsilon$ whenever both n and m are greater than N .

That this condition is necessary, i.e., the sequence converges only if $|S_n - S_m| < \varepsilon$ when n and m are greater than N , follows from the definition of convergence. Suppose that the sequence converges to a limit S . Then, for any positive ε' , there must be an integer N such that $|S_n - S| < \varepsilon'$ and $|S - S_m| < \varepsilon'$, for every n and m greater than N . By addition, $|S_n - S| + |S - S_m| < 2\varepsilon'$. Since

$$|S_n - S_m| = |(S_n - S) + (S - S_m)| \leq |S_n - S| + |S - S_m|,$$

it follows that $|S_n - S_m| < 2\varepsilon'$. Now, for any ε that might be chosen for the Cauchy test, we might choose $\varepsilon' = \varepsilon/2$. Then $|S_n - S_m| < \varepsilon$.

In order to prove the sufficiency of the Cauchy condition we must show (1) that the terms do not increase numerically beyond all bounds; (2) that the presence of infinitely many terms that lie between two finite bounds insures the existence of at least one limit, a limit being a number L such that there are infinitely many terms of the sequence between $L - \varepsilon$ and $L + \varepsilon$, for every positive ε ; and, finally, (3) that only one limit is possible. We consider these requirements in order.

(1) If the terms were to increase numerically beyond all bounds, the following situation would surely arise: for any positive M_1 and any subscript N there is a definite term $S_n, n > N$, such that $S_n > M_1$; now, for a second number, $M_2 = S_n + 100$, there is a definite term $S_m, m > N$, such that $S_m > M_2$; hence, $|S_m - S_n| > 100$. That is, the Cauchy condition cannot be satisfied.

(2) This is the critical point, the proof of which depends upon a deeper study of real numbers than has been supplied up to this time. (See Courant, Vol. I, p. 60.) At this juncture it may be considered an axiom concerning the real number system.

* Augustin-Louis Cauchy was one of the greatest of French mathematicians. It was perhaps due to his influence more than to that of any other scholar that calculus was put on a rigorous basis, early in the nineteenth century.

(3) Suppose the terms of the sequence have two limits in the sense of (2), say S and S' . Then, if we choose $\epsilon = \frac{1}{2}|S - S'|$, for sufficiently large n we shall have $|S - S_n| < \epsilon$ and $|S_n - S'| < \epsilon$. But this is impossible, since we cannot at the same time have $|S - S'| = 4\epsilon$ and $|S - S'| = |S - S_n + S_n - S'| \leq |S - S_n| + |S_n - S'| < 2\epsilon$.

The Cauchy test, as applied to series, becomes the following: the infinite series $a_1 + a_2 + a_3 + \cdots$ converges to a definite sum if, and only if, for each positive ϵ there exists an integer N such that $|a_n + a_{n+1} + \cdots + a_{n+k}| < \epsilon$, for every $n \geq N$ and every positive integer k . That is, after one passes the subscript N , the sum of any number of additional terms is less than ϵ .

The second theorem we shall state has to do with bounded, monotonic sequences. A sequence is bounded if there exists a positive number M such that $|S_n| < M$ for every subscript n , i.e., every term has its absolute value less than M . It is monotonic nondecreasing if $S_{n+1} \geq S_n$ for every subscript n ; and monotonic nonincreasing if $S_{n+1} \leq S_n$ for every subscript n . On the one hand, the terms never decrease as we proceed in the sequence, and on the other, they never increase.

THEOREM ON BOUNDED MONOTONIC NONDECREASING SEQUENCES. If the terms of a sequence satisfy the inequalities

$$S_1 \leq S_2 \leq S_3 \leq \cdots \leq S_n \leq \cdots \leq M,$$

the sequence has a limit which is not greater than M .

A similar theorem for bounded monotonic nonincreasing sequences insures a limit not less than a lower bound.

To prove this theorem, first divide the interval from S_1 to M into two equal parts, S_1R and RM , illustrated geometrically in Fig. 1. Now,

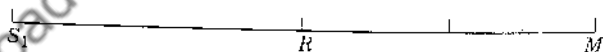


FIG. 1

choose the part containing infinitely many terms of the sequence. There is only one such part. Why? Divide this part again into equal parts, and again choose the one containing infinitely many terms of the sequence. This process can be continued indefinitely. Thus, if a positive ϵ were preassigned, the above process could be continued until infinitely many terms of the sequence fall in an interval of length less than ϵ , and only a finite number, say N , fall outside this interval. In this way, an integer N is found such that $|S_n - S_m| < \epsilon$ if n and m are both greater than N . Hence, by the Cauchy test, the sequence converges to a limit not greater than M .

An important application of this theorem to infinite series will be made in Art. 5.

EXERCISES

1. Prove that the sequence

$$S_1 = 1 + 1/1!, \quad S_2 = 1 + 1/1! + 1/2!, \quad S_3 = 1 + 1/1! + 1/2! + 1/3! \\ \dots\dots\dots S_n = 1 + 1/1! + 1/2! + \dots + 1/n!, \\ \dots\dots\dots$$

converges, by first showing that $S_n \leq 1 + 1 + 1/2 + 1/2^2 + \dots + 1/2^{n-1} < 3$.

2. By use of the inequality in Exercise 1 and the binomial formula, prove that the sequence, $S_n = (1 + 1/n)^n$, has a limit. (The limits of the sequences in Exercise 1 and Exercise 2 are the same, and equal to e . Compare Art. 5, Chapter 11.)

3. Show that the sequence $S_n = 1000 - 10^n/n!$ converges.

5. Fundamental Convergence Tests for Series of Positive Terms. The most fundamental test for convergence or divergence of an infinite series is the definition of the sum of an infinite series. In order to test a series by this means we must form the sequence of partial sums and test the convergence of this sequence. Thus, the burden is merely shifted from one convergence problem to another. This is, nevertheless, the final test to which all others must be referred. We proved the convergence of a geometric series with $r < 1$ by this means in Art. 3. In that article, also, the divergence of the harmonic series, $1 + \frac{1}{2} + \frac{1}{3} + \dots$, was proved by reference to the sequence of partial sums. The latter problem anticipated the comparison test for divergence, below.

Next to the definition, the most fundamental test for convergence or divergence is that of comparison with a series of known behavior. This test applies to series of positive terms. Series containing both positive and negative terms are more troublesome except in special cases, one of which is discussed in Art. 9. For series of positive terms, we state the following two comparison tests.

CONVERGENCE. Let it be required to test the series

$$a_1 + a_2 + a_3 + \dots + a_n + \dots, \quad (A)$$

in which all terms are positive. If a second series of positive terms,

$$c_1 + c_2 + c_3 + \dots + c_n + \dots, \quad (C)$$

can be found which is known to converge, and such that for some value of the subscript n onward, each term of series A is not greater than the corresponding term of the convergent series C , then series A converges to a sum that is not greater than the sum of series C .

DIVERGENCE. Let it be required to test the series

$$b_1 + b_2 + b_3 + \dots + b_n + \dots, \quad (B)$$

in which all terms are positive. If a second series of positive terms

$$d_1 + d_2 + d_3 + \dots + d_n + \dots, \quad (D)$$

can be found which is known to diverge, and such that for some value of the subscript n onward, each term of series B is not less than the corresponding term of the divergent series D , then series B diverges.

Since the deletion of a finite number of terms at the beginning of a series does not affect its convergence or divergence, these tests would be equally general in their application if we were to require that $a_n \leq c_n$ in the convergence test, and $b_n \geq d_n$ in the divergence test, for all values of n , i.e., from the start. If, for example, the inequality $a_n \leq c_n$ is satisfied only for $n > p$, we can delete the first p terms of both (A) and (C) and compare the series which remain.

In order to prove the convergence test, let us denote the partial sums of series (A) by S_n , and those of series (C) by S'_n . Since the terms of both of these series are positive, both S_n and S'_n increase with n . Let us suppose, in accordance with the above comment, that $a_n \leq b_n$ for all values of n . It follows that $S_n \leq S'_n$. Now, the sequence S'_n converges to a limit S' , and, moreover, $S'_n \leq S'$. Hence, $S_n \leq S'$. Thus, S_n is a sequence of nondecreasing terms, none of which is greater than S' . Therefore, by the theorem on bounded, monotonic sequences, S_n converges to a limit which is not greater than S' .

The proof of the divergence test is easier, and, in fact, has already been outlined in the discussion of the harmonic series in Art. 3. In general, we note that each partial sum of the series B is not less than the corresponding term of a sequence that increases beyond all bounds. Hence, the sequence of partial sums belonging to sequence B must also increase beyond all bounds.

EXAMPLE 1. Test the series

$$\frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots + \frac{1}{n!} + \cdots$$

for convergence or divergence.

SOLUTION. Comparing this series with the convergent geometric series,

$$1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^n} + \cdots,$$

we see that

$$\frac{1}{1!} = 1, \quad \frac{1}{2!} = \frac{1}{2}, \quad \frac{1}{3!} < \frac{1}{2^2}, \quad \frac{1}{4!} < \frac{1}{2^3}, \quad \cdots, \quad \frac{1}{n!} < \frac{1}{2^n}, \quad \cdots.$$

Hence, the series converges.

EXAMPLE 2. Test the series

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} + \cdots$$

for convergence.

SOLUTION. Comparing this series with the harmonic series

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots,$$

which was shown to diverge, we see that

$$\frac{1}{\sqrt{1}} = \frac{1}{1}, \quad \frac{1}{\sqrt{2}} > \frac{1}{2}, \quad \frac{1}{\sqrt{3}} > \frac{1}{3}, \quad \cdots, \quad \frac{1}{\sqrt{n}} > \frac{1}{n}, \quad \cdots.$$

Hence, the series diverges.

EXAMPLE 3. Test the series

$$\frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p} + \cdots,$$

which we shall call a *p-series*, for convergence, (a) if $p \leq 1$; (b) if $p > 1$.

SOLUTION. If $p = 1$, the *p-series* is identical with the harmonic series, $\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$, and hence diverges. If $p < 1$, each term of the *p-series* after the first is greater than the corresponding term of the harmonic series, and therefore the series diverges. If $p > 1$, we can show that the series converges, by a method similar to that used to show the divergence of the harmonic series, in Art. 3. We compare the *p-series* with the series

$$\frac{1}{1^p} + \left[\frac{1}{2^p} + \frac{1}{2^p} \right] + \left[\frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p} \right] + \cdots \quad (E)$$

Each term of the *p-series* is seen to be less than or equal to the corresponding term of series *E*. Collecting the terms of series *E* as indicated by the brackets, and noting that

$$\frac{1}{2^p} + \frac{1}{2^p} = \frac{2}{2^p} = \frac{1}{2^{p-1}};$$

$$\frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p} = \frac{4}{4^p} = \frac{1}{4^{p-1}} = \frac{1}{2^{2(p-1)}} = \left(\frac{1}{2^{p-1}} \right)^2;$$

etc., we see that series *E* is equivalent to the series

$$1 + \frac{1}{2^{p-1}} + \left(\frac{1}{2^{p-1}} \right)^2 + \left(\frac{1}{2^{p-1}} \right)^3 + \cdots.$$

This is a geometric series, with common ratio $= \frac{1}{2^{p-1}} < 1$. It follows that the geometric series converges, and, hence, by comparison, so does the *p-series* with $p > 1$.

The two most useful series for proving convergence by comparison are the geometric series with $r < 1$ and the p -series with $p > 1$. The most useful series for proving divergence by comparison is the harmonic series. For this purpose any harmonic series can be used (see Exercise 1).

In the use of the comparison test, one should apply the comparison to the general, or n th term, in order to be assured that the conclusion reached holds for all terms.

EXAMPLE 4. Test the series

$$\frac{1}{2 \cdot 3} + \frac{2}{3 \cdot 4} + \frac{3}{4 \cdot 5} + \cdots + \frac{n}{(n+1)(n+2)} + \cdots$$

for convergence.

SOLUTION. We note that the general term

$$\frac{n}{(n+1)(n+2)} > \frac{\frac{1}{2}(n+1)}{(n+1)(n+2)} \quad \text{if } n > 1;$$

i.e.,

$$\frac{n}{(n+1)(n+2)} > \frac{1}{2} \cdot \frac{1}{n+2}.$$

Now, the series having the general term $\frac{1}{2} \cdot \frac{1}{n+2}$ is

$$\frac{1}{2} \left(\frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{n+2} + \cdots \right).$$

Since the series in parentheses is a harmonic series, it diverges. Hence, the given series diverges.

EXERCISES

1. Prove that any harmonic series diverges. (A harmonic series is one whose terms are the reciprocals of an arithmetic series.)

Determine the convergence or divergence of each of the series in Exercises 2-15, by the comparison test.

2. $\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \cdots + \frac{1}{2n} + \cdots$

3. $1 + \frac{1}{2^2} + \frac{1}{3^3} + \frac{1}{4^4} + \cdots + \frac{1}{n^n} + \cdots$

4. $\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n(n+1)} + \cdots$

$$5. 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt[3]{3}} + \cdots + \frac{1}{\sqrt[n]{n}} + \cdots.$$

$$6. \frac{2}{1 \cdot 3} + \frac{4}{3 \cdot 5} + \frac{6}{5 \cdot 7} + \cdots + \frac{2n}{(2n-1)(2n+1)} + \cdots.$$

$$7. \frac{3}{2 \cdot 3 \cdot 4} + \frac{5}{3 \cdot 4 \cdot 5} + \frac{7}{4 \cdot 5 \cdot 6} + \cdots + \frac{2n+1}{(n+1)(n+2)(n+3)} + \cdots.$$

$$8. \frac{1}{2^{1/2}} + \frac{1}{2} + \frac{1}{2^{3/2}} + \cdots + \frac{1}{2^{n/2}} + \cdots.$$

$$9. \frac{1}{\sqrt{1 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 4}} + \cdots + \frac{1}{\sqrt{n(n+1)}} + \cdots.$$

$$10. \frac{1}{3} + \frac{1}{5} + \frac{1}{9} + \cdots + \frac{1}{2^n + 1} + \cdots.$$

$$11. \frac{1}{\log 2} + \frac{1}{\log 3} + \frac{1}{\log 4} + \cdots + \frac{1}{\log(n+1)} + \cdots.$$

$$12. \frac{1}{1+1^3} + \frac{2}{1+2^3} + \frac{3}{1+3^3} + \cdots + \frac{n}{1+n^3} + \cdots.$$

$$13. \frac{2}{2^2} + \frac{4}{3^2} + \frac{6}{4^2} + \cdots + \frac{2n}{(n+1)^2} + \cdots.$$

$$14. \frac{1}{2} + \frac{1 \cdot 2}{3 \cdot 4} + \frac{1 \cdot 2 \cdot 3}{4 \cdot 5 \cdot 6} + \cdots + \frac{(n!)^2}{(2n)!} + \cdots.$$

$$15. \frac{1}{3} + \frac{1 \cdot 2}{3 \cdot 5} + \frac{1 \cdot 2 \cdot 3}{3 \cdot 5 \cdot 7} + \cdots + \frac{(n!)^2 2^n}{(2n+1)!} + \cdots.$$

$$16. \text{ If } f(n) = \frac{2n}{2n+1} \text{ and } g(n) = \frac{2n-1}{2n}, \text{ which of the following series converge?}$$

$$(a) \sum_{i=1}^{\infty} f(i).$$

$$(b) \sum_{i=1}^{\infty} f(i) \cdot g(i).$$

$$(c) \sum_{i=1}^{\infty} [f(i) - g(i)].$$

$$(d) \sum_{i=1}^{\infty} \frac{f(i)}{g(i)}.$$

6. Comparison with an Integral. The facts concerning the convergence of the p -series suggest an analogy with the improper integral $\int_1^{\infty} \frac{1}{x^p} dx$. This integral converges if $p > 1$, and diverges otherwise. A test for the convergence of an improper integral over an infinite range was based on this result (Chapter 17).

Suppose that some positive function $f(x)$ is given, which tends monotonically to 0 as x tends to ∞ , in such a way that

$$\int_1^{\infty} f(x) dx \quad (6.1)$$

converges. We can show that the infinite series

$$\sum_{i=1}^{\infty} f(i) \quad (6.2)$$

must converge, also. If, on the other hand, integral (6.1) diverges, then series (6.2) also diverges. In order to make this theorem plausible, we

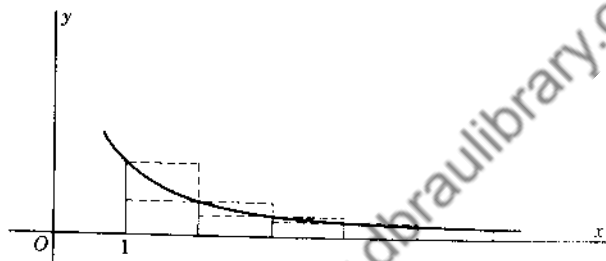


FIG. 2

consider the infinitely long area bounded by the ordinate $x = 1$, the x -axis, and the curve $y = f(x)$ (Fig. 2). The series

$$f(1) + f(2) + f(3) + \cdots \quad (6.2)$$

gives the sum of the areas of the rectangles extending above the curve in the figure. Clearly, this sum is greater than the integral, and, thus, if the integral diverges, so does the series.

The sum of the areas of the rectangles that lie below the curve in the figure is

$$f(2) + f(3) + f(4) + \cdots \quad (6.3)$$

Hence, if the integral converges, so does series 6.3. But series (6.2) differs from series (6.3) only by the single term $f(1)$, and, hence, series (6.2) also converges.

EXAMPLE 1. Test the series

$$\frac{1}{2 \log 2} + \frac{1}{3 \log 3} + \frac{1}{4 \log 4} + \cdots$$

for convergence.

SOLUTION. This series is $\sum_1^{\infty} f(i)$, where $f(i) = \frac{1}{(i+1) \log(i+1)}$. Hence, we shall examine the integral

$$\begin{aligned} \int_1^{\infty} \frac{1}{(x+1) \log(x+1)} dx &= \lim_{A \rightarrow \infty} \int_1^A \frac{1}{(x+1) \log(x+1)} dx \\ &= \lim_{A \rightarrow \infty} \left[\log \log(x+1) \right]_1^A \\ &= \lim_{A \rightarrow \infty} [\log \log(A+1) - \log \log 2]. \end{aligned}$$

As $A \rightarrow \infty$, so does $\log(A+1)$ and $\log \log(A+1)$. Hence, the integral diverges, and the series also diverges.

EXAMPLE 2. Test the series

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots$$

for convergence, by the integral test.

SOLUTION. This series is $\sum_1^{\infty} f(i)$, where $f(i) = \frac{1}{i(i+1)}$. Hence, the integral to be used is $\int_1^{\infty} \frac{1}{x(x+1)} dx$. The convergence of this integral follows from Test III, Art. 5, Chapter 17, since $1/x(x+1) < 1/x^2$.

This example is identical with Exercise 4, Art. 5, where the series was to be shown convergent by the comparison test.

EXERCISES

Test the convergence of the series in Exercises 1-5 by comparison with an integral.

1. $\frac{1}{2} + \frac{2}{5} + \frac{3}{10} + \cdots + \frac{n}{n^2+1} + \cdots$

2. $\frac{1}{2} + \frac{1}{5} + \frac{1}{10} + \cdots + \frac{1}{n^2+1} + \cdots$

3. $\frac{1}{2} \div \frac{2}{17} \div \frac{3}{82} + \cdots + \frac{n}{n^4+1} + \cdots$

4. $\frac{1}{1 \cdot 3} + \frac{1}{2 \cdot 4} + \frac{1}{3 \cdot 5} + \cdots$

5. $\frac{1}{2+2 \log 2} + \frac{1}{3+3 \log 3} + \frac{1}{4+4 \log 4} + \cdots$

6. Prove that the p -series converges if $p > 1$ and diverges if $p \leq 1$, by comparison with an integral.

7. Show that the series $\sum_1^{\infty} \frac{1}{(i+1)[\log(i+1)]^p}$ converges if $p > 1$, and diverges if $p \leq 1$.

8. Make use of the fact that $\lim_{A \rightarrow \infty} \frac{A}{e^A} = 0$ to prove the convergence of

$$(a) \frac{1}{e} + \frac{2}{e^2} + \frac{3}{e^3} + \cdots \quad (b) \frac{1}{e} + \frac{4}{e^2} + \frac{9}{e^3} + \cdots \quad (c) \frac{1}{e} + \frac{8}{e^2} + \frac{27}{e^3} + \cdots$$

9. Generalize the results of Exercise 8.

10. Find an upper bound to the sum of each of the convergent series. (a) $\sum_{1}^{\infty} \frac{1}{i^2}$
(b) $\sum_{1}^{\infty} \frac{1}{i^{3/2}}$ (c) $\sum_{1}^{\infty} \frac{1}{i^{1.1}}$

7. A Ratio Test and a Root Test. Comparison with a geometric series leads to the following test for convergence of series of positive terms.

TEST. The series $a_1 + a_2 + a_3 + \cdots + a_n + \cdots$ converges if, for all terms a_n from some term onward, say a_k , the inequality

$$a_n < br^{n-1} \quad (8.1)$$

holds, where b and r are fixed positive numbers with $r < 1$.

Several other inequalities may be used in place of (8.1), any one of which may be more convenient in a given case than (8.1). Two of these are alternative forms of a ratio test; two others are alternative forms of a root test.

RATIO TEST. From the infinite series

$$a_1 + a_2 + a_3 + \cdots$$

form the general expression for the value of the ratio of any term to the preceding term, viz., $\frac{a_{n+1}}{a_n}$. The series converges if, from some term onward,

$$\frac{a_{n+1}}{a_n} < r \quad (8.2)$$

where r is a positive number less than 1. On the other hand, if, from some term onward,

$$\frac{a_{n+1}}{a_n} \geq 1 \quad (8.3)$$

the series diverges.

Alternatively, if

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = L < 1 \quad (8.4)$$

the series converges. If

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} > 1 \quad (8.5)$$

the series diverges, whereas if this limit is exactly 1, the test is inconclusive.

It is not hard to show that if (8.4) holds, so does (8.2); and if (8.2) holds, so does (8.1). We leave the proof to the student, in Exercise 16. The converse is not valid in either case.

ROOT TEST. The infinite series

$$a_1 + a_2 + a_3 + \cdots$$

converges if, from some term onward,

$$\sqrt[n]{a_n} < r \quad (8.6)$$

where r is a positive number less than 1. On the other hand, if, from some term onward,

$$\sqrt[n]{a_n} \geq 1 \quad (8.7)$$

the series diverges.

Alternatively, if

$$\lim_{n \rightarrow \infty} \sqrt[n]{a_n} = L < 1 \quad (8.8)$$

the series converges. If

$$\lim_{n \rightarrow \infty} \sqrt[n]{a_n} > 1 \quad (8.9)$$

the series diverges. If this limit is exactly 1, the test is inconclusive.

As with the ratio test, if (8.8) holds, so does (8.6); and if (8.6) holds, so does (8.1). The proof is left to the student, Exercise 17. The converse is not valid in either case.

It is important to observe that in both (8.2) and (8.6), the quantity on the left side of the inequality sign must be less than some fixed positive number which is, itself, less than 1. Neither the inequality $\frac{a_{n+1}}{a_n} < 1$ nor $\sqrt[n]{a_n} < 1$ insures convergence. For example, both inequalities hold for the harmonic series.

In applying the root test, two results from Chapter 11, Art. 14 (or Chapter 20, Art. 8) are useful: $\lim_{n \rightarrow \infty} \sqrt[n]{a} = 1$, if a is any positive constant, and $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.

EXAMPLE 1. Test the series

$$\frac{1}{10} + \frac{1 \cdot 2}{10^2} + \frac{1 \cdot 2 \cdot 3}{10^3} + \cdots + \frac{n!}{10^n} + \cdots$$

by the ratio test.

SOLUTION. Here $a_{n+1} = \frac{(n+1)!}{10^{n+1}}$ and $a_n = \frac{n!}{10^n}$. Hence,

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \lim_{n \rightarrow \infty} \left[\frac{(n+1)!}{10^{n+1}} \bigg/ \frac{n!}{10^n} \right] = \lim_{n \rightarrow \infty} \frac{n+1}{10} = \infty.$$

The series therefore diverges.

EXAMPLE 2. Test the series

$$\frac{1}{3} + \left(\frac{2}{5}\right)^2 + \left(\frac{3}{7}\right)^3 + \cdots + \left(\frac{n}{2n+1}\right)^n + \cdots$$

by the root test.

SOLUTION. We have

$$\begin{aligned}\lim_{n \rightarrow \infty} \sqrt[n]{a_n} &= \lim_{n \rightarrow \infty} \sqrt[n]{\left(\frac{n}{2n+1}\right)^n} \\ &= \lim_{n \rightarrow \infty} \frac{n}{2n+1} = \frac{1}{2} < 1.\end{aligned}$$

The series therefore converges.

EXERCISES

Use the ratio test or the root test to determine the convergence or divergence of each of the series in Exercises 1-15.

- $\frac{3}{2} + \frac{4}{2^2} + \frac{5}{2^3} + \frac{6}{2^4} + \cdots$
- $\frac{7}{6} + \frac{7^2}{2 \cdot 6^2} + \frac{7^3}{3 \cdot 6^3} + \frac{7^4}{4 \cdot 6^4} + \cdots$
- $10 + \frac{10^2}{2!} + \frac{10^3}{3!} + \frac{10^4}{4!} + \cdots$
- $\frac{1}{1} + \frac{1 \cdot 2}{1 \cdot 3} + \frac{1 \cdot 2 \cdot 3}{1 \cdot 3 \cdot 5} + \cdots + \frac{n!}{1 \cdot 3 \cdot 5 \cdots (2n-1)} + \cdots$
- $\frac{0}{1 \cdot 2} + \frac{1}{2 \cdot 2^2} + \frac{2}{3 \cdot 2^3} + \cdots + \frac{n-1}{n \cdot 2^n} + \cdots$
- $.9 + 2(.9)^2 + 3(.9)^3 + \cdots + n(.9)^n + \cdots$
- $\frac{2}{1 \cdot 2} + \frac{2^2}{2 \cdot 3} + \frac{2^3}{3 \cdot 4} + \cdots + \frac{2^n}{n(n+1)} + \cdots$
- $\frac{1}{2} + \left(\frac{2}{3}\right)^4 + \left(\frac{3}{4}\right)^9 + \cdots + \left(\frac{n}{n+1}\right)^{n^4} + \cdots$
- $1 + \frac{3^4}{2^4 \cdot 2^2} + \frac{4^9}{3^9 \cdot 2^3} + \cdots + \frac{(n+1)^{n^4}}{n^{n^4} \cdot 2^n} + \cdots$
- $\frac{1}{3} + \frac{1 \cdot 3}{3 \cdot 6} + \frac{1 \cdot 3 \cdot 5}{3 \cdot 6 \cdot 9} + \cdots + \frac{1 \cdot 3 \cdots (2n-1)}{3 \cdot 6 \cdots 3n} + \cdots$
- $\frac{1}{\log 2} + \frac{1}{2(\log 2)^2} + \frac{1}{3(\log 2)^3} + \cdots + \frac{1}{n(\log 2)^n} + \cdots$

$$12. \frac{1}{\log 3} + \frac{4}{(\log 3)^2} + \frac{9}{(\log 3)^3} + \cdots + \frac{n^2}{(\log 3)^n} + \cdots$$

$$13. \frac{2}{1} (1 - \varepsilon) + \frac{2 \cdot 4}{1 \cdot 3} (1 - \varepsilon)^2 + \frac{2 \cdot 4 \cdot 6}{1 \cdot 3 \cdot 5} (1 - \varepsilon)^3 + \cdots \\ + \frac{2 \cdot 4 \cdot 6 \cdots 2n}{1 \cdot 3 \cdot 5 \cdots (2n - 1)} (1 - \varepsilon)^n + \cdots, \quad (0 < \varepsilon < 1).$$

$$14. 1 + \frac{2^2}{2!} + \frac{3^3}{3!} + \cdots + \frac{n^n}{n!} + \cdots$$

$$15. \frac{2}{1.1} - \frac{(3/2)^2}{(1.1)^2} + \frac{(4/3)^3}{(1.1)^3} + \cdots + \frac{(1 + 1/n)^n}{(1.1)^n} + \cdots$$

16. Prove the ratio test in form (8.2). [Hint: If $\frac{a_{n+1}}{a_n} < r$, then $a_{n+1} < a_n r$; $a_{n+2} < a_{n+1} r < a_n r^2$, etc.]

17. Prove the root test in form (8.6). [Hint: If $\sqrt[n]{a_n} < r$, then $a_n < r^n$.]

18. Prove the ratio test and the root test for divergence.

8. Absolute Convergence. Thus far we have dealt almost exclusively with series of positive terms. If both positive and negative terms are present, convergence may result from a tendency of terms of opposite sign to cancel one another, rather than from the way in which succeeding terms tend to 0. For example, the harmonic series diverges; but if the signs are made alternately positive and negative, the resulting series, $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$, converges, as will be shown in Art. 9. We distinguish between series, on the one hand, whose convergence depends upon the arrangement of positive and negative signs, and those, on the other hand, whose convergence results from the size of the terms, by calling the first of these *conditionally* convergent, and the second *absolutely* convergent.

DEFINITION. An absolutely convergent series is one such that the series formed from the absolute values of the terms converges.

For example, $1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \cdots$ is absolutely convergent, since

$$|1| + |-\frac{1}{2}| + |\frac{1}{4}| + |-\frac{1}{8}| + |\frac{1}{16}| + \cdots$$

or

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$$

converges. Similarly,

$$1 + \frac{1}{2^2} - \frac{1}{3^2} + \frac{1}{4^2} - \frac{1}{5^2} + \frac{1}{6^2} - \frac{1}{7^2} + \frac{1}{8^2} + \cdots$$

converges absolutely. If each term is replaced by its absolute value, we have the familiar p -series with $p = 2$.

On the other hand, although the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

will be shown to converge, it does not converge absolutely, since the series of absolute values is the harmonic series, which diverges.

The following theorem, which we shall not prove, is intuitively reasonable.

THEOREM. If a series converges absolutely, it also converges in the usual sense. That is, if $|a_1| + |a_2| + |a_3| + \cdots$ converges, so does $a_1 + a_2 + a_3 + \cdots$.

Any of the tests discussed in Arts. 5, 6, and 7, for convergence of series of positive terms, can be applied without change to series of absolute values and thus can be used to determine absolute convergence.

9. Alternating Series. Although we shall not enter into a discussion of conditionally convergent series in general, one type of series of positive and negative terms deserves special attention. This is an alternating series, in which the signs of consecutive terms alternate. The test for the convergence of such a series is especially easy.

TEST FOR ALTERNATING SERIES. An alternating series,

$$a_1 - a_2 + a_3 - a_4 + \cdots,$$

in which all the a 's are positive, is convergent if, from some term onward, $a_{n+1} \leq a_n$ and $\lim_{n \rightarrow \infty} a_n = 0$.

This theorem can be proved as follows. First, without loss of generality, we may assume that the absolute value of each term is less than that of its predecessor, from the start. We now write the sum of an even number of terms in two ways.

$$S_{2n} = (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2n-1} - a_{2n}), \quad (9.1)$$

and

$$S_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n}. \quad (9.2)$$

From (9.1) we see that S_{2n} increases as n increases, since each difference in parentheses is positive or 0. From (9.2), for a similar reason, we see that S_{2n} is never greater than a_1 . Hence, from the theorem on bounded, monotonic sequences, S_{2n} tends to a limit S as n tends to infinity. Since S_{2n} is always the sum of an even number of terms, we must examine the sums containing an odd number of terms, to see if these sums tend to the same limit. Now, the odd sum, $S_{2n+1} = S_{2n} + a_{2n+1}$, and since $a_{2n+1} \rightarrow 0$ as $n \rightarrow \infty$,

$$\lim_{n \rightarrow \infty} S_{2n+1} = \lim_{n \rightarrow \infty} S_{2n} + \lim_{n \rightarrow \infty} a_{2n+1} = S.$$

Thus, the whole sequence of partial sums tends to the limit S .

An additional result, that gives a ready estimate of the error when the sum of a convergent alternating series is approximated by a finite number of terms, is obtained from (9.1) and an expression for S_{2n+1} corresponding to (9.2).

$$S_{2n+1} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n} - a_{2n+1}). \quad (9.3)$$

From (9.1) we see that S_{2n} increases with n so that $S_{2n} \leq S$; whereas from (9.3) we see that S_{2n+1} decreases as n increases, so that $S \leq S_{2n+1}$. Subtracting S_{2n} from both sides of the latter inequality, we have

$$(S - S_{2n}) \leq (S_{2n+1} - S_{2n}) = a_{2n+1}.$$

Subtracting S and S_{2n} , successively, from S_{2n-1} , and noting that $S \geq S_{2n}$, we have

$$(S_{2n-1} - S) \leq (S_{2n-1} - S_{2n}) = a_{2n}.$$

These results mean that any partial sum differs from the sum of the whole series by a quantity that is not more than the absolute value of the first term following the partial sum. That is, any partial sum approximates the sum of the infinite series, with an error not greater than the first term omitted.

EXAMPLE. Calculate the sum of the series

$$.1 - \frac{(.1)^2}{2} + \frac{(.1)^3}{3} - \frac{(.1)^4}{4} + \frac{(.1)^5}{5} + \cdots$$

to 4 decimal places.

SOLUTION. Since the general term of the series is $\pm \frac{(.1)^n}{n}$, which tends to 0 as n tends to ∞ , the series converges. A partial sum will have 4-place accuracy if the first term omitted is less than .00005. We see that the fourth term has the value .000025, and hence we keep 3 terms, obtaining

$$\begin{aligned} .1 - .005 + .00033 &= .09533 \\ &= .0953, \text{ to 4 places.} \end{aligned}$$

Compare this series, and the result, with the series for $\log 1.1$ in Art. 4, Chapter 20.

EXERCISES

Determine which of the series in Exercises 1-8 converge, and which converge absolutely.

$$1. 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

$$2. 1 - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \cdots$$

3. $1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \cdots$
4. $\frac{1}{1 \cdot 2} - \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} - \frac{1}{4 \cdot 5} + \cdots$
5. $\frac{1}{\log 2} - \frac{1}{\log 3} + \frac{1}{\log 4} - \frac{1}{\log 5} + \cdots$
6. $\sum_{i=1}^{\infty} \frac{1}{(2i-1) \cos[(i-1)\pi]}$
7. $\sum_{i=1}^{\infty} (-1)^{i-1}/i^{\frac{1}{2}}$
8. $\sum_{i=1}^{\infty} (-1)^i(i+1)/2i$
9. $\sum_{i=2}^{\infty} (-1)^{i-1}(\log i)^{-\frac{1}{2}}$
10. $\sum_{i=1}^{\infty} (-1)^{i-1}/(2+i^{-1})$

11. What is the maximum error incurred by taking the first 3 terms to represent the sum of the series in the case of each convergent series in Exercises 1-10?

12. Calculate the sum of each of the following series correct to 3 decimal places:

- (a) $1 - \frac{1}{2^4} + \frac{1}{3^4} - \frac{1}{4^4} + \cdots$
- (b) $1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots$
- (c) $\frac{1}{2} - \frac{1}{3}(\frac{1}{2})^3 + \frac{1}{5}(\frac{1}{2})^5 - \frac{1}{7}(\frac{1}{2})^7 + \cdots$
- (d) $\frac{\log 2}{10} - \frac{\log 3}{10^2} + \frac{\log 4}{10^3} - \frac{\log 5}{10^4} + \cdots$

10. The Representation of Numbers by Infinite Series. A repeating decimal is really a concise form of a certain infinite geometric series. Thus,

$$.3333 \cdots = .3 + .03 + .003 + .0003 + \cdots$$

This is a geometric series with common ratio equal to .1. Its sum is

$$S = \frac{a}{1-r} = \frac{.3}{1-.1} = \frac{1}{3}$$

Similarly, the repeating decimal

$$.142857142857 \cdots = .142857 + .000000142857 + \cdots$$

This is a geometric series with common ratio equal to .000001. Its sum is

$$S = \frac{a}{1-r} = \frac{.142857}{1-.000001} = \frac{.142857}{.999999} = \frac{1}{7}$$

Every terminating decimal can be expressed in another way by means of a repeating decimal. Thus,

$$.13 = .129999 \dots$$

This is easily verified by writing

$$\begin{aligned} .129999 \dots &= .12 + .009 + .0009 + .00009 + \dots \\ &= .12 + \text{a geometric series.} \end{aligned}$$

Since the series

$$.009 + .0009 + .00009 + \dots = \frac{.009}{1-.1} = \frac{.009}{.9} = .01,$$

we see that $.129999 \dots = .12 + .01 = .13$.

The preceding examples show how rational numbers can be represented by infinite series. The representation of irrational numbers by infinite series is of greater interest, since a convergent-series representation of such a number provides a means of approximating it to any desired degree of accuracy. At this point, we give some examples without proof. These representations will be justified by the results of Chapter 22. They are also reminiscent of Chapter 20.

$$1. e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots \quad (\text{Exercise 1, Art. 6, Chapter 20}).$$

$$2. \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \quad (\text{Example 7, Art. 2, Chapter 20}).$$

$$3. \log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \quad (\text{Exercise 3, Art. 4, Chapter 20}).$$

$$\begin{aligned} 4. \sqrt{6} &= 2 \left[1 + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2!} \cdot \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)^2 \right. \\ &\quad \left. + \frac{1}{3!} \cdot \frac{1}{2} \cdot \left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot \left(\frac{1}{2}\right)^3 + \dots \right] \\ &= 2 + \frac{1}{2} - \frac{1}{16} + \frac{1}{64} + \dots \quad (\text{Example 2, Art. 7, Chapter 20}). \end{aligned}$$

$$5. \sin(1 \text{ rad.}) = 1 - \frac{1}{3!} + \frac{1}{5!} - \frac{1}{7!} + \dots \quad (\text{Exercise 1, Art. 7, Chapter 20}).$$

11. Summary. This chapter was concerned with the convergence or divergence of infinite series of constants. We first defined the sum of an infinite series. Proof that a given series converges must show that there is a number satisfying this definition. All tests rest basically on this definition. There are two problems here, however: (1) proof that there is a sum, and (2) determination of this sum. Our attention was given almost exclusively to the first of these problems. The second is less important, and, in fact, is often no problem at all, since the expressed infinite sum may be taken as the symbol of the sum, just as $a + b$ both indicates an addition and denotes the sum of that addition. It is often desirable, however, to express this sum in a more concise form. This problem is taken up in many textbooks devoting more space to the subject of series, under the heading of "Summation of Series."*

Several tests for convergence were given. Except for the special test for alternating series, all of them apply to series of positive terms, and hence to the question of absolute convergence of any series. The latter question has decisive bearing at many points in the further development of series. This fact will appear in some of the discussions in Chapter 22.

MISCELLANEOUS EXERCISES

Test for convergence by any means. Examine alternating series for absolute convergence.

$$1. 1 + \frac{2}{1.1} + \frac{3}{(1.1)^2} + \frac{4}{(1.1)^3} + \cdots$$

$$2. 1 - \frac{2}{\log 2} + \frac{3}{\log 3} - \frac{4}{\log 4} + \cdots$$

$$3. 1 - \frac{1}{\log \sqrt{2}} + \frac{1}{\log \sqrt{3}} - \frac{1}{\log \sqrt{4}} + \cdots$$

$$4. \frac{1}{2} + \frac{\sqrt{2}}{3} + \frac{\sqrt{3}}{4} + \frac{\sqrt{4}}{5} + \cdots$$

$$5. \frac{1}{\sqrt{2}} - \frac{2}{\sqrt{3}} + \frac{3}{\sqrt{10}} - \frac{4}{\sqrt{17}} + \cdots$$

$$6. (\sqrt{2} - 1) + (\sqrt{3} - 2) + \cdots + (\sqrt{n^2 + 1} - n) + \cdots$$

$$7. \sqrt{2} + (\sqrt{17} - \sqrt{15}) + \cdots + (\sqrt{n^4 + 1} - \sqrt{n^4 - 1}) + \cdots$$

* Knopp, K.: *Theory and Applications of Infinite Series*, Chapter VIII (Advanced, theoretical), and Hall and Knight: *Higher Algebra*, Chapter 29.

$$8. \sum_{i=1}^{\infty} \frac{i^a}{i!}, \text{ where } a = \text{any positive integer or fraction.}$$

$$9. \sum_{i=1}^{\infty} \left[\left(\frac{i+1}{i} \right)^{i+1} - \frac{i+1}{i} \right]^{-i}.$$

$$10. \sum_{i=1}^{\infty} \frac{i}{1 + i\sqrt{i+1}}.$$

$$11. \frac{3}{4} - \frac{3 \cdot 4}{4 \cdot 6} + \frac{3 \cdot 4 \cdot 5}{4 \cdot 6 \cdot 8} - \frac{3 \cdot 4 \cdot 5 \cdot 6}{4 \cdot 6 \cdot 8 \cdot 10} + \dots.$$

$$12. \sum_{i=0}^{\infty} 1/(a + \sqrt[i]{i}), \quad a > 0.$$

$$13. \sum_{i=0}^{\infty} i/(a + ib), \quad a > 0, b > 0.$$

$$14. \sum_{i=2}^{\infty} 1/(i^2 - i).$$

$$15. \sum_{i=1}^{\infty} 1/(i\sqrt[i]{i}).$$

Downloaded from www.dbraulibrary.org.in

Representation of Functions by Power Series

1. The Possibility of Expressing a Function by a Series. In the last chapter we saw how a series of constant terms, even though infinite in number, might converge to a finite quantity. If these terms were variable, in particular, if all were functions of the same variable x , it is possible that the series might converge for all values of x in some interval. Now, we recall that if a rule, formula, or set of directions of any sort is given, by which some definite number is found to correspond to each value of a variable x on some range, a function of x is thereby defined on this range. An infinite series of functions of x , all defined on a common interval, converging for all values of x on this interval, would constitute such a rule, and, hence, would define a function of x on the interval.

It may be asked how one would be led to an infinite series of functions converging in some interval. Such a series may arise in many ways, perhaps the most important of which is the construction of a function in the form of an infinite series to satisfy specified requirements, e.g., to satisfy the differential equation $d^2y/dx^2 + xy = 0$ and have $y = 12$ and $dy/dx = 4$ when $x = 0$. This particular problem is solved as Example 2, Art. 8.

The main problem of this chapter is that of determining series representations of functions that are also known in another form. It is natural to ask why we should want to do this. What can be the advantage if substituting for a known form anything so unwieldy as the sum of infinitely many terms? The answer lies in the fact that this unwieldiness is more apparent than real. Actually, a sum of even infinitely many terms, combined, as they are, by the simplest operation of addition, is more manageable than many concise expressions involving the more complicated operations of division, root-taking, calculation of logarithms, etc. Such advantages as may be gained by the avoidance of division and root-taking, for example, of course require that the individual terms of a series be free of these operations. Hence, it is to be expected that the simplest type of series will be one formed by the addition of positive integral powers of a variable x , multiplied by constant coefficients, constituting a *power series*.

We shall find, in Art. 2, that if such a series of powers of x converges for a value of x different from 0, it converges for every value in an interval. Thus, it defines a function in this interval. It will be found that many operations can be performed upon functions defined by series of this sort, valid within an interval of convergence, by applying these operations to the series in the same way that we apply them to the simplest of algebraic functions which they resemble, viz., polynomials. This is true for addition, subtraction, and multiplication. Moreover, derivatives and integrals of functions represented by such series are found by termwise differentiation and integration of the series, exactly as these processes are carried out with polynomials.

2. The Interval of Convergence of a Power Series. The question of convergence of a series of functions of a variable goes beyond that of convergence of a series of constants. The question now is: for what values of the variable, if any, does the series converge? As the variable changes, the terms change, and the series may change from a convergent to a divergent one. We define the *interval of convergence* of a power series as follows.

DEFINITION. The interval of convergence of a power series in x is the interval of values such that the series converges for every value of x in this interval and diverges for all others.

We shall consider some examples.

EXAMPLE 1. Find the interval of convergence of the series

$$1 + x + x^2 + x^3 + \cdots$$

SOLUTION. For $|x| < 1$, this series is geometric with $|r| < 1$. Hence, it converges absolutely for $|x| < 1$. It diverges if $|x| > 1$. If $x = 1$, the series is also clearly divergent. If $x = -1$, the series becomes $1 - 1 + 1 - 1 + \cdots$. Since the sequence of partial sums oscillates between 1 and 0, the series does not converge. The interval of convergence is thus $|x| < 1$, or $-1 < x < 1$.

EXAMPLE 2. The series $1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$ converges absolutely for all values of x . This is shown by the ratio test. We have

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\left| \frac{x^{n+1}}{(n+1)!} \right|}{\left| \frac{x^n}{n!} \right|} = \left| \frac{x}{n+1} \right|.$$

For any value of x , the fraction $\frac{x}{n+1} \rightarrow 0$ as $n \rightarrow \infty$, and, hence, the series converges for every x .

EXAMPLE 3. The series $1 + x + 2!x^2 + 3!x^3 + \cdots$ converges for $x = 0$, and for no other values. Even if x has a very small value, after a

sufficient number of terms have been passed, the terms increase with n , and so do not tend to 0. Thus, if $x = .001$, and $n = 1000$,

$$n!x^n = 1 \cdot 2 \cdot \cdots 1000 \cdot (.001)^{1000},$$

and the next term in the series is

$$1 \cdot 2 \cdot \cdots 1000 \cdot 1001 \cdot (.001)^{1001} = 1000! (.001)^{1000} \cdot 1001 (.001).$$

Since $1001 (.001) > 1$, we see that the 1001th term is greater than the 1000th term. Thereafter, the terms increase more and more rapidly.

EXAMPLE 4. Investigate the convergence of the series

$$x - x^2/2 + x^3/3 - x^4/4 + \cdots$$

SOLUTION. We first find the range in which the series is absolutely convergent. For this purpose we examine the ratio

$$\frac{|a_{n+1}|}{|a_n|} = \frac{|x^{n+1}|}{|n+1|} \bigg/ \frac{|x^n|}{|n|} = \frac{n}{n+1} |x|.$$

Since $\lim_{n \rightarrow \infty} \frac{n}{n+1} = 1$, we have $\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} = |x|$. Hence the series is absolutely convergent if $|x| < 1$. If $x = 1$, the series is alternating, with $|a_n| = 1/n$, and hence converges. But, if $x = -1$, the series becomes $-1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \cdots$, which diverges. Thus, the series converges absolutely if $|x| < 1$, and its whole interval of convergence is $-1 < x \leq 1$.

From these examples, it appears that the interval of convergence of a series of powers of x always centers around $x = 0$. This is reflected in the following theorem.

THEOREM. If an infinite series of powers of x , $\sum_{i=1}^{\infty} c_i x^i$, converges for $x = b$, it converges absolutely for every number numerically smaller than b , i.e., for $|x| < |b|$.

To prove this theorem, we note that $c_i b^i \rightarrow 0$ as $i \rightarrow \infty$, which is a necessary condition for the convergence of the series $\sum_{i=1}^{\infty} c_i b^i$. We can therefore find a bound M , such that $|c_i b^i| < M$ for all values of i . Now, if $|x| < |b|$, we have

$$|c_i x^i| = |c_i b^i| \cdot \frac{|x|^i}{|b|^i} < M \frac{|x|^i}{|b|^i}.$$

It follows that the series $\sum_{i=1}^{\infty} c_i x^i$ converges, by comparison with the convergent geometrical series $\sum_{i=1}^{\infty} M \frac{|x|^i}{|b|^i}$.

Notice that the interval given by this theorem is open. The series may or may not converge for $x = -b$. As a consequence of the theorem, a series converges absolutely throughout every open interval within the interval of convergence of the series. To determine the complete interval of convergence we can first find the values for which the series converges absolutely, as in Example 4, and then test the end-points of the interval thus found.

A more general form of a power series than those so far given is a series of powers of $(x - k)$, in which k can be any constant. Conclusions about power series in x are valid with little change for series of this type. In fact, by replacing $x - k$ by u , such a series is transformed into one of the simpler type. The interval of convergence for a power series in $x - k$ must center around $x = k$, which corresponds to $u = 0$.

EXERCISES

Find the interval of convergence of each of the following series.

- $1 + 2x + 3x^2 + 4x^3 + \cdots$
- $1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$
- $x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots$
- $x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$
- $1 - x + x^2 - x^3 + \cdots$
- $\frac{x}{1 \cdot 2} + \frac{x^2}{2 \cdot 2^2} + \frac{x^3}{3 \cdot 2^3} + \cdots$
- $x + \frac{1}{2} \cdot \frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \cdot \frac{x^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \cdot \frac{x^7}{7} + \cdots$
- $1 - \frac{1}{2} \cdot \left(\frac{x}{2}\right)^2 + \frac{1 \cdot 3}{2 \cdot 4} \cdot \left(\frac{x}{2}\right)^4 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \cdot \left(\frac{x}{2}\right)^6 + \cdots$
- $(x - 2) - \frac{(x - 2)^3}{\sqrt{2}} + \frac{(x - 2)^5}{\sqrt{3}} - \frac{(x - 2)^7}{\sqrt{4}} + \cdots$
- $\sum_{i=1}^{\infty} i(x + 1)^i$
- $\sum_{i=1}^{\infty} \frac{(3x)^i}{i^2}$
- $\sum_{i=1}^{\infty} \left(\frac{x - a}{a}\right)^{i-1}$
- $\sum_{i=1}^{\infty} \frac{i}{5^i} \left(\frac{x}{2}\right)^{i-1}$
- $\sum_{i=1}^{\infty} i^2 3^i (2x - 1)^i$
- $\sum_{i=1}^{\infty} (\cos in) \frac{(x - 1)^i}{i}$

3. Examples of Power Series for Known Functions. It is easy to write power series for the functions discussed in Chapter 20, Art. 2. If we imagine the division continued indefinitely instead of terminating with a remainder, we obtain

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + \cdots, \quad (3.1)$$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + \cdots, \quad (3.2)$$

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n} + \cdots. \quad (3.3)$$

The validity of these equations is still to be tested, of course. Obviously, none of them is valid for $|x| \geq 1$, since all three diverge for such values. From our work in Chapter 20 we know that they are all valid in the range $|x| < 1$. The sum of series (3.1), for example, is

$$S = \lim_{n \rightarrow \infty} S_n, \text{ in which } S_n = 1 + x + x^2 + \cdots + x^{n-1}.$$

But, from Chapter 20, we know that

$$\begin{aligned} \frac{1}{1-x} &= 1 + x + x^2 + \cdots + x^{n-1} + R_n(x) \\ &= S_n + R_n \end{aligned}$$

and that $R_n \rightarrow 0$ as $n \rightarrow \infty$. That is,

$$\lim_{n \rightarrow \infty} S_n = S \quad \text{and} \quad \lim_{n \rightarrow \infty} S_n = \frac{1}{1-x}$$

and, hence, $S = 1/(1-x)$.

In the same way, we can show that the series in (3.2) and (3.3) converge to $1/(1+x)$ and $1/(1+x^2)$, respectively, for $|x| < 1$.

Formulas for $\log(1+x)$, $\log(1-x)$, $\log(1+x)/(1-x)$, and $\tan^{-1}x$ from Chapter 20, are also readily extended to power series representations of these functions. Thus,

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n+1} \frac{x^n}{n} + \cdots$$

valid for $-1 < x \leq 1$;

$$\log(1-x) = -\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^n}{n} + \cdots\right)$$

valid for $-1 \leq x < 1$;

$$\log \frac{1+x}{1-x} = 2 \left(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{2n-1}}{2n-1} + \cdots \right)$$

valid for $-1 < x < 1$;

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + \cdots$$

valid for $-1 \leq x \leq 1$.

Proof, in each of these cases, that the series converges to the function that appears on the left side of the equality sign, is the same as for (3.1). It should be carefully noted that the proof consists in showing that $R_n(x)$, which is the difference between the given function and the first n terms of the series, converges to 0; and not in simply showing that the series converges. It is possible for a series to converge, but to something different from the function from which it is obtained by some formal rule.

A distinction between the equations in this article and those in Chapter 20 lies in the fact that, unlike the above equations, the validity of an equation that expresses a function as the sum of a polynomial and a remainder R_n does not depend upon the convergence of R_n to 0. For example

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^{n-1} + \frac{x^n}{1-x}$$

holds for all values of x except $x = 1$, for which it has no meaning; but

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \cdots$$

holds only for those values for which $\frac{x^n}{1-x} \rightarrow 0$ as $n \rightarrow \infty$.

EXERCISES

1. Find a power-series representation of the function $1/(1-2x)$. Within what interval is it valid?
2. Find a power-series representation of the function $1/(2-x)$. Within what interval is it valid? [Hint: Write $1/(2-x) = \frac{1}{2} \cdot 1/(1-\frac{1}{2}x)$.]
3. Find a power-series representation of the function $1/(a+bx)$. Within what interval is it valid. [See hint in Exercise 2.]
4. Find a power-series representation of the function $1/(1-x^2)$ in two ways: (1) in the same way as in the example $1/(1+x^2)$ above, and (2) by splitting $1/(1-x^2)$ into partial fractions, obtaining a series for each, and combining the results. What can be said of the validity of the second method?

5. The terms of the power series for $\log(1+x)$ are equal, respectively, to integrals of corresponding terms of the power series for $1/(1+x)$. Can we conclude, in general, that if

$$f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n + \cdots,$$

$$g(x) = \int_0^x f(t) dt = a_0 x + \frac{1}{2} a_1 x^2 + \frac{1}{3} a_2 x^3 + \cdots + \frac{1}{n+1} a_n x^{n+1} + \cdots?$$

Justify your answer. If the answer is negative, must we conclude that our derivation of series 3.1 for $\log(1+x)$ is invalid? Give reasons for your answer.

6. Write infinite series for (a) $\log 3$. (b) $\pi/4$.

7. By using two series discussed in this article, show that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \frac{2}{3} - \frac{2}{3 \cdot 3^3} + \frac{2}{5 \cdot 3^5} - \frac{2}{7 \cdot 3^7} + \cdots$$

$$\frac{3}{5} - \frac{3^3}{3 \cdot 5^3} + \frac{3^5}{5 \cdot 5^5} - \frac{3^7}{7 \cdot 5^7} + \cdots$$

8. Write $\log(2-x)$ as a power series in x , and find the interval of convergence.

9. Write $\log(a+bx)$ as a power series in x , and find the interval of convergence.

10. Write $\tan^{-1}(x-1)$ as a power series in $(x-1)$ and find the interval of convergence.

4. Taylor's Series. If $f(x)$ possesses continuous derivatives of all orders at $x = a$, we can form the series

$$f(a) + f'(a)(x-a) + \frac{1}{2!} f''(a)(x-a)^2 + \cdots$$

$$+ \frac{1}{n!} f^{(n)}(a)(x-a)^n + \cdots \quad (4.1)$$

This is a *Taylor's series*, corresponding to $f(x)$. It results from extending the polynomials given by Taylor's theorem to infinitely many terms. Note that we do not equate $f(x)$ with this series, even though the series is obtained from this function. For, series (4.1) may not converge to $f(x)$; in fact, it may not converge at all for some values of x . Or it may converge to values different from those of $f(x)$. Of course, we are interested in the series mainly in the interval, if any, in which it converges to $f(x)$. This is determined solely by the convergence to 0 of $R_n(x)$, the remainder in Taylor's theorem. Within the interval in which $R_n(x) \rightarrow 0$ as $n \rightarrow \infty$ we may put $f(x)$ equal to series 4.1.

The special form assumed by (4.1) when $a = 0$, is

$$f(0) + f'(0)x + \frac{1}{2!} f''(0)x^2 + \cdots + \frac{1}{n!} f^{(n)}(0)x^n + \cdots \quad (4.2)$$

Series 4.2 is seen to be a power series in x and has wide application. It is often called Maclaurin's series, although there seems to be no historical justification for this designation.

It is clear that Taylor's series is constructed by the same process by which we have already (in Chapter 20) formed approximating polynomials by Taylor's theorem. In setting a given function equal to its Taylor's series, we must recognize that the validity of the equality is limited to the interval in which the series converges to the function.

EXAMPLE. Write the Taylor's series for $\sin 2x$ in powers of x , and find the range in which the series is equal to the function.

SOLUTION.

$$f(x) = \sin 2x \qquad f(0) = 0$$

$$f'(x) = 2 \cos 2x \qquad f'(0) = 2$$

$$f''(x) = -4 \sin 2x \qquad f''(0) = 0$$

$$f'''(x) = -8 \cos 2x \qquad f'''(0) = -8,$$

etc. From the alternating recurrence of sines and cosines, the cyclic arrangement of signs, and the repeated multiplication by 2, we see that the series is

$$\begin{aligned} \sin 2x = 2x - \frac{8}{3!}x^3 + \frac{32}{5!}x^5 - \frac{256}{7!}x^7 + \cdots \\ + (-1)^{n-1} \frac{2^{2n-1}}{(2n-1)!} x^{2n-1} + \cdots \end{aligned}$$

The student can show that the equality is valid for all values of x .

EXERCISES

Find the Taylor's series in powers of x for the functions in Exercises 1-10, including the general term, and find the interval in which the series converges to the function.

$$1. e^x, \quad 2. \sin x, \quad 3. \cos x, \quad 4. \sinh x, \quad 5. \cosh x.$$

$$6. e^{-2x}, \quad 7. \log(1 + 2x), \quad 8. \sqrt{1+x}, \quad 9. \frac{1}{\sqrt{1-x}}, \quad 10. xe^x.$$

Obtain 4 terms of the Taylor's series in powers of x , for the functions in Exercises 11-15.

$$11. \log(1 + e^x), \quad 12. e^{\sin x}, \quad 13. \log(1 + \cos x).$$

$$14. e^x \sin x, \quad 15. (1-x) \cdot \log(1-x).$$

Obtain 4 terms of the Taylor's series in powers of $x-2$, for the functions in Exercises 16-20.

$$16. \log x, \quad 17. e^x, \quad 18. e^{x+1}, \quad 19. \cosh x, \quad 20. \sqrt{2+x}.$$

21. By substituting $x^2 = u$ in the expansion for e^{-u} in powers of u , find an expansion of e^{-x^2} .
22. Obtain a power series for $x \cos x$ from that for $\cos x$.
23. Obtain a power series for $(\sin x)/x$ from that for $\sin x$.
24. Substitute $u^2 = x$ in the expansion of $1/\sqrt{1-u}$ in Exercise 9. Hence, find an expansion for $\sin^{-1} u$. Are all steps valid? Explain.
25. Use the method of Exercise 24 to obtain a series for $\sinh^{-1} u$.

5. Uniform Convergence of Power series. Suppose the power series, $\sum c_i x^i = c_0 + c_1 x + c_2 x^2 + \cdots$ converges absolutely for $|x| < M$ (or for all values of x). That is, $|c_0| + |c_1 x| + |c_2 x^2| + \cdots$ converges if $|x| < M$. If we write

$$\sum_0^\infty |c_i x^i| = |c_0| + |c_1 x| + |c_2 x^2| + \cdots = R_n,$$

where $R_n = \sum_n^\infty |c_i x^i|$, it must follow that $R_n \rightarrow 0$ as $n \rightarrow \infty$. More precisely, if x is specified, for any positive ε there must be a positive integer N such that, for $n > N$, $\sum_n^\infty |c_i x^i| < \varepsilon$. Now, let x lie in the interval $|x| \leq p$, where p is some number less than M . Each term in $\sum_n^\infty |c_i x^i|$ will assume its maximum value in this interval when $x = p$. Hence, if we find the integer N such that when $n > N$, $\sum_n^\infty |c_i p^i| < \varepsilon$, it must follow that $\sum_n^\infty |c_i x^i| < \varepsilon$ for all values of x in the interval $|x| \leq p$. Thus, the function to which the series converges in the interval $|x| \leq p$ can be approximated as closely as we please *throughout the interval* by a polynomial of n terms, i.e., n can be taken once for all large enough so that $R_n < \varepsilon$, no matter what value of x is used. Because we can meet any standard of approximation, throughout the interval, by one uniform choice of the number of terms to be used, the series is said to *converge uniformly* in this interval.

The above demonstration proves the following theorem.

THEOREM. If a series of powers of x converges absolutely for $x < M$, the convergence is uniform in any closed interval $|x| \leq p$, where $p < M$.

The series $1 + x + x^2 + \cdots$ converges uniformly to $1/(1-x)$ in the interval $|x| < \frac{1}{2}$. The fact that the error could be made less than $\frac{1}{4}$ throughout this interval by using 3 terms was pointed out in Chapter 20, Art. 3. At the same time, although the series converges for all values of x in the open interval $|x| < 1$, it does not converge uniformly in this

interval, since, as x is taken closer to 1, more and more terms are required to make R_n small.

It was shown in Art. 2 that a series of powers of x that converges for $x = b$ converges absolutely for $|x| < |b|$. Hence, the series converges absolutely and uniformly in any interval $|x| \leq p < |b|$.

The property of uniform convergence is of the greatest importance in the study of infinite sequences and series of functions. It should be realized that its introduction with reference to power series has a special simplicity. It is largely because of the property of uniform convergence enjoyed by power series that they are so useful and manageable.

6. Theorems on Power Series. The fact that the convergence of a power series is absolute and uniform within an interval (excepting the trivial case of series which converge only at a single point, e.g., $x = 0$) leads to several important theorems about functions defined by power series.

THEOREM 1. (CONTINUITY). A power series represents a continuous function $f(x)$ within the interval of convergence. The word "within" here means "in the interior of," exclusive of end-points.

We have to prove that $\lim_{h \rightarrow 0} f(x+h) = f(x)$, if x lies within the interval of convergence, i.e., $|f(x+h) - f(x)|$ can be made less than any chosen positive ε by taking h sufficiently small. For this purpose, we write

$$f(x) = P_n(x) + R_n(x), \quad f(x+h) = P_n(x+h) + R_n(x+h),$$

where $P_n(x)$ is the polynomial consisting of the first n terms of the series. Now,

$$\begin{aligned} |f(x+h) - f(x)| &= |P_n(x+h) - P_n(x) + R_n(x+h) - R_n(x)| \\ &\leq |P_n(x+h) - P_n(x)| + |R_n(x+h)| + |R_n(x)|. \end{aligned}$$

Since x is within the interval of convergence, there is an interval around x throughout which the series converges uniformly. For all sufficiently small h , $x+h$ will be in this interval, and we can take n so great that both $|R_n(x+h)|$ and $|R_n(x)|$ are less than $\varepsilon/3$. Since $P_n(x)$ is a polynomial, it is continuous, and, hence, we can take h small enough so that $|P_n(x+h) - P_n(x)| < \varepsilon/3$. Therefore,

$$|f(x+h) - f(x)| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon,$$

and the continuity of $f(x)$ is proved.

THEOREM 2. (INTEGRATION). If $\sum_{0}^{\infty} c_i x^i$ converges to the function $f(x)$ within an interval, then

$$\int_a^b f(x) dx = \int_a^b c_0 dx + \int_a^b c_1 x dx + \int_a^b c_2 x^2 dx + \cdots,$$

if both a and b lie in the interval of convergence. That is, the infinite series of integrals of the terms of the power series converges to the integral of $f(x)$.

To prove this, we write

$$f(x) = \sum_0^n c_i x^i + R_n(x).$$

Hence, integrating the polynomial $\sum_0^n c_i x^i$, term by term,

$$\int_a^b f(x) dx = \int_a^b c_0 dx + \int_a^b c_1 x dx + \cdots + \int_a^b R_n(x) dx,$$

and

$$\left| \int_a^b f(x) dx - \sum_0^n \int_a^b (c_i x^i) dx \right| = \left| \int_a^b R_n(x) dx \right|.$$

On account of the uniform convergence of the series $\sum_0^\infty c_i x^i$, n can be chosen so large that $|R_n(x)|$ is less than any preassigned positive ε , for all values of x in the interval of integration. Thus,

$$\left| \int_a^b f(x) dx - \sum_0^n \int_a^b (c_i x^i) dx \right| < \left| \int_a^b \varepsilon dx \right| = \varepsilon |b - a|.$$

Since $\varepsilon |b - a|$ can be made as small as we please, we have

$$\begin{aligned} \int_a^b f(x) dx &= \lim_{n \rightarrow \infty} \sum_0^n \int_a^b (c_i x^i) dx \\ &= \sum_0^\infty \int_a^b (c_i x^i) dx. \end{aligned}$$

It follows as a corollary to this theorem that

$$\int_a^x f(t) dt = \sum_0^\infty \int_a^x (c_i t^i) dt,$$

if a and x lie within the interval of convergence of $\sum_0^\infty c_i x^i$. In particular, if $a = 0$,

$$\begin{aligned} \int_0^x f(t) dt &= \sum_0^\infty \int_0^x c_i t^i dt \\ &= c_0 x + \frac{1}{2} c_1 x^2 + \frac{1}{3} c_2 x^3 + \cdots + \frac{1}{n} c_{n-1} x^n + \cdots \quad (6.1) \end{aligned}$$

Now, $\int_0^x f(t) dt$ is an inverse derivative of $f(x)$, and the terms in (6.1) are inverse derivatives of the terms of the given series. Hence, if a power

series $\sum_0^{\infty} c_i x^i$ converges to a function $f(x)$, the power series of inverse derivatives converges to an inverse derivative of $f(x)$. That is,

$$\int f(x) dx = C + \sum_{i=0}^{\infty} c_i x^{i+1}/(i+1)$$

or

$$\int \left[\sum_{i=0}^{\infty} c_i x^i \right] dx = \sum_{i=0}^{\infty} \left[\int c_i x^i dx \right] = C + \sum_{i=0}^{\infty} c_i x^{i+1}/(i+1).$$

THEOREM 3. (DIFFERENTIATION). If $\sum_{i=0}^{\infty} c_i x^i$ converges to the function $f(x)$ within an interval, then the series of derivatives, $\sum_{i=0}^{\infty} \frac{d}{dx} (c_i x^i)$, converges to the derivative of $f(x)$, within the same interval; i.e., $f'(x) = \sum_{i=0}^{\infty} c_i i x^{i-1}$.

To prove this theorem we first show that the series of derivatives has the same interval of convergence as the given series. Applying the ratio test to both series, we have for the given series,

$$\frac{|a_{n+1}|}{|a_n|} = \frac{|c_{n+1} x^{n+1}|}{|c_n x^n|} = \left| \frac{c_{n+1}}{c_n} \right| \cdot |x|,$$

giving the interval of convergence $|x| < \lim_{n \rightarrow \infty} |c_n/c_{n+1}|$. For the derivative series,

$$\frac{|a_{n+1}|}{|a_n|} = \frac{|(n+1)c_{n+1} x^n|}{|nc_n x^{n-1}|} = \frac{n+1}{n} \left| \frac{c_{n+1}}{c_n} \right| |x|.$$

Hence, this series also converges for

$$|x| < \lim_{n \rightarrow \infty} \left[\frac{n}{n+1} \left| \frac{c_n}{c_{n+1}} \right| \right] = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right|.$$

It now follows that the derivative series converges uniformly within the interval of convergence of the given series to some function $g(x)$. Hence, by Theorem 2, this series can be integrated term by term to give

the series $\sum_0^{\infty} c_i x^i$, which must be equal to $\int g(x) dx$. Thus, since

$$\sum_0^{\infty} c_i x^i = f(x) = \int g(x) dx,$$

it follows that

$$f'(x) = g(x) = \sum_0^{\infty} c_i i x^{i-1}.$$

Theorems 2 and 3 can be summarized briefly: a power series can be integrated and differentiated term by term, within the interval of convergence. It follows naturally that termwise integration or differentiation can be performed as many times as we wish. We make use of this right of repeated differentiation to prove the following theorem.

THEOREM 4. (UNIQUENESS). A function has no more than one series representation in powers of x .

Suppose $f(x)$ has two such representations in the same interval $|x| < A$,

$$f(x) = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n + \cdots,$$

and

$$f(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n + \cdots.$$

Subtracting,

$$0 = (b_0 - c_0) + (b_1 - c_1)x + (b_2 - c_2)x^2 + \cdots + (b_n - c_n)x^n + \cdots \quad (6.2)$$

Putting $x = 0$, we have $b_0 - c_0 = 0$, or $b_0 = c_0$. Differentiating (6.2),

$$0 = (b_1 - c_1) + 2(b_2 - c_2)x + \cdots + n(b_n - c_n)x^{n-1} + \cdots.$$

Again putting $x = 0$, $b_1 = c_1$. Differentiating again, and putting $x = 0$ in the result, we get $b_2 = c_2$. By a continuation of the process, we conclude that $b_0 = c_0$, $b_1 = c_1$, $b_2 = c_2$, $\cdots$, $b_n = c_n$, $\cdots$. Hence, the two series representations are identical.

As an immediate corollary to this theorem, we can state that a series representation in powers of x , no matter how obtained, is the Taylor's series of the function in the Maclaurin form.

The extension of these four theorems to powers of $x - k$, in place of x , is immediate. As regards Theorem 4, however, although a function is expressible in only one way by a series of powers of $x - k$, it might be expressible by a series in $x - k_1$ and a different series in $x - k_2$. If the intervals of convergence of these two series should overlap, there would be points at which the function could be expressed by different power series, but in powers of different quantities, $x - k_1$ and $x - k_2$, respectively.

7. Power Series and Polynomials. The theorems of Art. 6 establish several properties that convergent power series enjoy in common with polynomials. It is easy to show, in addition, that a polynomial can be added to or subtracted from a power series, and that a power series can be multiplied or divided by a polynomial. Further, two power series, converging in the same interval, can be added, subtracted, multiplied, or divided. Let us consider addition and subtraction. Suppose

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots,$$

$$g(x) = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n + \cdots,$$

both series converging in the same interval. By adding the series term by term, we obtain a third power series

$$(a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_n + b_n)x^n + \cdots \quad (7.1)$$

It is easy to see that this series converges to $f(x) + g(x)$. For, we have

$$f(x) = \lim_{n \rightarrow \infty} S_n, \quad \text{where } S_n = a_0 + a_1x + \cdots + a_nx^n,$$

and

$$g(x) = \lim_{n \rightarrow \infty} S'_n, \quad \text{where } S'_n = b_0 + b_1x + \cdots + b_nx^n.$$

Now,

$$S_n + S'_n = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n,$$

which is the sum of n terms of (7.1). Hence, the sum of (7.1) is

$$\lim_{n \rightarrow \infty} (S_n + S'_n) = \lim_{n \rightarrow \infty} S_n + \lim_{n \rightarrow \infty} S'_n = f(x) + g(x).$$

In the same way, we find that the infinite series

$$(a_0 - b_0) + (a_1 - b_1)x + (a_2 - b_2)x^2 + \cdots + (a_n - b_n)x^n + \cdots$$

has the sum:

$$\lim_{n \rightarrow \infty} (S_n - S'_n) = \lim_{n \rightarrow \infty} S_n - \lim_{n \rightarrow \infty} S'_n = f(x) - g(x).$$

Termwise multiplication is also valid in a common interval of convergence, although we shall not prove this fact. The procedures for dealing with division will also be omitted.

EXAMPLE 1. Find 5 terms of the power series in x for the function $e^x \log(1+x)$.

SOLUTION. We have

$$\begin{aligned} e^x \log(x+1) &= (1+x+x^2/2!+x^3/3!+x^4/4!+x^5/5!+x^6/6!+\cdots) \\ &\quad \cdot (x-x^2/2+x^3/3-x^4/4+x^5/5-x^6/6+\cdots) \\ &= x+x^2/2+x^3/3+3x^5/40-7x^6/144+\cdots \end{aligned}$$

The product series must converge in the interval common to the intervals of convergence of the factors. Since the series for e^x converges everywhere, and that for $\log(1+x)$ converges in the interval $-1 < x \leq 1$, the product series also converges for $-1 < x \leq 1$.

EXAMPLE 2. Obtain the power series in x for $(1+x) \cdot \log(1+x)$ by integrating that for $\log(1+x)$.

SOLUTION. Since $\int \log(1+x) dx = (1+x) \log(1+x) - x$, we only have to integrate the series for $\log(1+x)$ and add x . Thus,

$$(1+x) \log(1+x)$$

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

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

The interval of convergence of this series is the same as that of the series for $\log(1+x)$, viz., $-1 < x \leq 1$.

EXERCISES

Obtain 5 terms of the power series in x for each of the following functions, and find the interval of convergence for each.

1. $\sin x + \cos x$.

2. $\sin x/(1-x)$.

3. $e^x \cos x$.

4. $(\sin x)/e^x$.

5. $x \log(x+2)$.

6. $[\log(1-x)]/(1+2x)$.

7. $xe^x - e^x = \int xe^x dx$.

8. $x \tan^{-1} x$.

9. $\int_0^x e^{-t^2} dt$.

10. $\int_0^x \frac{\sin t}{t} dt$.

8. Power Series Solutions of Differential Equations. Among the important applications of power series is that of the construction of a function in the form of a power series to satisfy a differential equation. We illustrate this use by means of examples. For each we assume the existence of a solution in this form. For the differential equations here encountered, this is guaranteed by an existence theorem in the theory of differential equations.

EXAMPLE 1. Find a function, in the form of a power series, that satisfies the differential equation $dy/dx = 2y$ and has the value 10 when $x = 0$.

SOLUTION. We assume that

$$y = c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots + c_nx^n + \cdots \quad (8.1)$$

Since $y = 10$ when $x = 0$, we know that $c_0 = 10$. Differentiating, which is permitted by Theorem 3, Art. 6,

$$dy/dx = c_1 + 2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1} + \cdots \quad (8.2)$$

Substituting from (8.2) and (8.1), with $c_0 = 10$, in the differential equation, we obtain

$$\begin{aligned} c_1 + 2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1} + \cdots \\ = 20 + 2c_1x + 2c_2x^2 + \cdots + 2c_nx^n + \cdots. \end{aligned}$$

Since these two series must represent identically the same function, their coefficients must be equal, by Theorem 4, Art. 6. Hence,

$$c_1 = 2 \cdot 10, \quad 2c_2 = 2c_1 = 2^2 \cdot 10, \quad 3c_3 = 2c_2 = \frac{2^3 \cdot 10}{2 \cdot 1}$$

$$c_2 = \frac{2^2 \cdot 10}{2 \cdot 1}, \quad c_3 = \frac{2^3 \cdot 10}{3 \cdot 2 \cdot 1},$$

$$\dots \dots \dots$$

$$nc_n = 2c_{n-1}.$$

The last result obtained is a *recursion formula* from which, with the values of c_0 to c_3 already found, we can write all coefficients. Thus, since

$4c_4 = 2c_3$, we have $c_4 = \frac{2^4 \cdot 10}{4 \cdot 3 \cdot 2 \cdot 1}$, etc. Hence,

$$\begin{aligned} y = 10 + 10 \cdot 2x + \frac{10 \cdot 2^2 x^2}{2!} + \frac{10 \cdot 2^3 x^3}{3!} + \frac{10 \cdot 2^4 x^4}{4!} + \cdots \\ + \frac{10 \cdot 2^n x^n}{n!} + \cdots \\ = 10 \left(1 + 2x + \frac{(2x)^2}{2!} + \frac{(2x)^3}{3!} + \cdots + \frac{(2x)^n}{n!} + \cdots \right). \end{aligned}$$

The series in parentheses is recognized as the power series for e^{2x} , and therefore $y = 10e^{2x}$. The solution could have been found in this form, as in Chapter 11.

EXAMPLE 2. Find 6 terms of a function, in the form of a power series in x , that satisfies the differential equation $d^2y/dx^2 + xy = 0$, and having $y = 12$ and $dy/dx = 4$ when $x = 0$.

SOLUTION. We write

$$y = c_0 + c_1x + c_2x^2 + c_3x^3 + c_4x^4 + \cdots + c_nx^n + \cdots \quad (8.3)$$

$$dy/dx = c_1 + 2c_2x + 3c_3x^2 + 4c_4x^3 + \cdots + nc_nx^{n-1} + \cdots \quad (8.4)$$

$$d^2y/dx^2 = 2c_2 + 3 \cdot 2c_3x + 4 \cdot 3c_4x^2 + \cdots + n(n-1)c_nx^{n-2} + \cdots \quad (8.5)$$

Since $y = 12$ and $dy/dx = 4$ when $x = 0$, we have, from (8.3) and (8.4): $c_0 = 12$, $c_1 = 4$. Substituting the series in (8.3) and (8.5) in the differential equation, we have

$$2c_2 + 3 \cdot 2c_3x + 4 \cdot 3c_4x^2 + \cdots + n(n-1)c_nx^{n-2} \\ + c_0x + c_1x^2 + c_2x^3 + c_3x^4 + \cdots + c_{n-3}x^{n-2} + c_{n-2}x^{n-1} + \cdots = 0,$$

or

$$2c_2 + (3 \cdot 2c_3 + c_0)x + (4 \cdot 3c_4 + c_1)x^2 + \cdots \\ + [n(n-1)c_n + c_{n-3}]x^{n-2} + \cdots = 0.$$

Since this power series is identically 0, the coefficients of the separate powers must be 0. Therefore, in addition to $c_0 = 12$, $c_1 = 4$, we have

$$2c_2 = 0, \quad c_2 = 0.$$

$$3 \cdot 2c_3 + c_0 = 0, \quad c_3 = -\frac{1}{3 \cdot 2} c_0 = -\frac{1}{3 \cdot 2} \cdot 12.$$

$$4 \cdot 3c_4 + c_1 = 0, \quad c_4 = -\frac{1}{4 \cdot 3} c_1 = -\frac{1}{4 \cdot 3} \cdot 4.$$

$$\dots \dots \dots$$

$$n(n-1)c_n + c_{n-3} = 0, \quad c_n = -\frac{1}{n(n-1)} c_{n-3}.$$

From the last result, which is a recursion formula giving each coefficient in terms of those preceding it, we can write as many terms as we wish. We have

$$c_0 = 12; \quad c_1 = 4; \quad c_2 = 0;$$

$$c_3 = -\frac{1}{3 \cdot 2} c_0 = -\frac{1}{3 \cdot 2} \cdot 12; \quad c_4 = -\frac{1}{4 \cdot 3} c_1 = -\frac{1}{4 \cdot 3} \cdot 4;$$

$$c_5 = -\frac{1}{5 \cdot 4} c_2 = 0; \quad c_6 = -\frac{1}{6 \cdot 5} c_3 = \frac{1}{6 \cdot 5 \cdot 3 \cdot 2} \cdot 12;$$

$$c_7 = -\frac{1}{7 \cdot 6} c_4 = -\frac{1}{7 \cdot 6 \cdot 4 \cdot 3} \cdot 4; \quad c_8 = -\frac{1}{8 \cdot 7} c_5 = 0, \text{ etc.}$$

Putting these values in (8.3), we have, for the first 6 terms:

$$y = 12 + 4x - 2x^3 - \frac{1}{3}x^4 + \frac{1}{15}x^6 + \frac{1}{126}x^7 - \dots$$

EXERCISES

In each of the following exercises, obtain at least 4 terms of a power series in x , satisfying the differential equation and the subsidiary conditions.

1. $\frac{dy}{dx} = y$; $y = 1$, when $x = 0$.
2. $(1+x)\frac{dy}{dx} + y = 0$; $y = 1$, when $x = 0$.
3. $(1+x)\frac{dy}{dx} = \frac{1}{2}y$; $y = 2$, when $x = 0$.
4. $(1+x)\frac{dy}{dx} = 2$; $y = -1$, when $x = 0$.
5. $\left(\frac{dy}{dx} - y\right)(1+x) = e^x$; $y = 5$, when $x = 0$. [Express e^x as a series.]
6. $\frac{d^2y}{dx^2} + y = 0$; $y = 2$, $\frac{dy}{dx} = 1$, when $x = 0$.
7. $\frac{d^2y}{dx^2} - y = 0$; $y = 2$, $\frac{dy}{dx} = 0$, when $x = 0$.
8. $\frac{d^2y}{dx^2} = y \cos x$; $y = 12$, $\frac{dy}{dx} = 6$, when $x = 0$.

9. Limitations of Power Series. It is implicit in our discussion of power series that a function can be represented in this way only if it possesses derivatives of all orders. The function $x^{\frac{1}{2}}$, for example, cannot be expressed by a series of powers of x , since the first derivative fails to exist at $x = 0$. Any fractional power of x presents a similar difficulty. So does $\log x$, $\cot x$, $1/(e^x - 1)$, and many other functions. For a similar reason, $(x-1)^{\frac{1}{2}}$, $\log(x-1)$, etc., cannot be expanded in powers of $x-1$.

The requirement of differentiability to all orders is really a stringent one. Many functions having practical applicability cannot meet this requirement. Differentiability to a lower order, say order n , permits the representation of a function by a polynomial of n terms and a remainder, by Taylor's theorem. This is often preferable to an infinite-series representation. Other series representations, e.g., in positive and negative powers, or in trigonometric terms, although posing difficulties not arising with positive integral powers, can be used for the representation of many functions that do not meet the strict requirements for power-series representation. The study of such series is a proper topic for a more advanced course.*

* Courant, Vol. 1, pp. 383-456; Further theory of power series; Trigonometric series. Churchill, R. V., *Fourier Series and Boundary Value Problems*; Application of trigonometric series to the solution of differential equations. Knopp, K., *Theory and Applications of Infinite Series*; General and thorough treatment of series.

Graphical Representation in Three Dimensions

1. The Place of Graphs in the Study of Functions. The study of functional relations is so closely bound to the graphical interpretations of such relations that the student may easily fall into the error of identifying the two. In this book, the operations of differentiation and integration were first developed in direct connection with their geometrical interpretations in rectangular coordinates. That these interpretations were not essential was frequently brought out, and this was given added emphasis by the introduction of different interpretations in polar coordinates.

Graphical representations proved so advantageous in the study of functions of one variable that it is natural to try to extend this use to the study of functions of two variables. Although we shall not be able to extend it further, many of the results obtained for functions of two variables can be easily generalized to functions of three or more variables without dependence upon graphical interpretations.

Our main problem is that of extending the operations of differentiation and integration to functions of two or more variables. In order to prepare the way for carrying out this purpose, we shall devote the present chapter to the discussion of the graphical representation of functions of two variables.

2. Rectangular Coordinates in Three Dimensions. In order to represent a functional relation of the sort $z = f(x, y)$, in a way similar to that used for graphical representation of $y = f(x)$, three perpendicular axes are needed. On an x -axis we lay off a value of the independent variable x , on a y -axis a value of the independent variable y , and the corresponding value of the function, or dependent variable z , is then measured on the third, or z -axis.

Of course, some distortion is inevitable in suggesting three mutually perpendicular axes on a flat sheet. Graphical representations of functions of two variables are, consequently, more difficult to visualize than are those of functions of one variable. The student must keep in mind the three-dimensional character of the figures, which is only suggested by two-dimensional drawings.

Figures 1, 2, and 3 show the mutually perpendicular axes. Figures 1 and 2 are alike, except for point of view. Figure 3 differs from Figs. 1 and 2 in the orientation of the axes. Figure 1 is called a *right-hand system*: if the x - and y -axes were the cross slots in a right-hand screw head, a rotation from the x -direction toward the y -direction would send the screw in the z -direction (Fig. 4). Figure 3 is called a *left-hand system*,

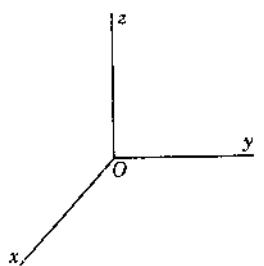


FIG. 1

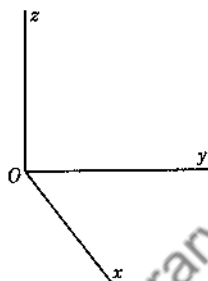


FIG. 2

for a similar reason. Although both systems are in common use, we shall employ the right-hand system exclusively. The whole system is represented in Fig. 5, the dotted lines representing the negative parts of the three axes.

We should try to visualize not only the three axes, but the three coordinate planes Oxy , Oxz , and Oyz . A distance along the z -axis is

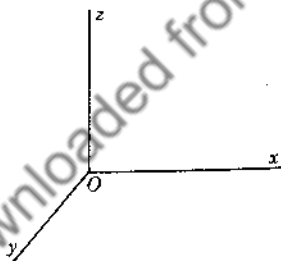


FIG. 3

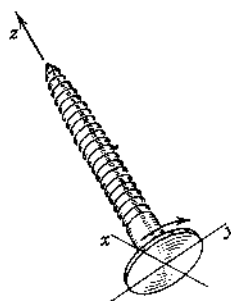


FIG. 4

to be thought of as a distance from the xy -plane, etc. These three mutually perpendicular planes divide space into eight parts, called *octants*, comparable to the four quadrants into which two coordinate axes divide the plane.

Each point in space has three coordinates, when referred to a given rectangular coordinate system. The signed, or directed, distance of a point from the yz -plane is its x -coordinate, that from the xz -plane its

y -coordinate, and that from the xy -plane its z -coordinate. It is conventional to write the coordinates of a point in the x, y, z order. Thus, point A in Fig. 5 is $(2, 5, 7)$. Similarly, points B and C are $(3, -2, -1)$ and $(0, 0, -4)$, respectively. If a point has unknown coordinates, we often refer to it as point (x, y, z) . The notation (x_1, y_1, z_1) usually refers to a fixed point, whereas (x, y, z) may denote any one of a collection of points.

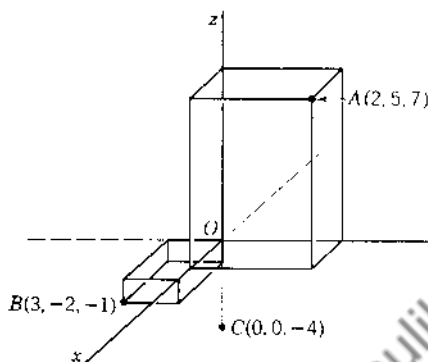


FIG. 5

3. Locus Problems in Three Dimensions. The fundamental problems of coordinate geometry of three dimensions are the same as in two dimensions, viz., (1) to find the locus of all points whose coordinates satisfy a given equation, and (2) to find the equation satisfied by the coordinates of the points of a given locus and by no other values.

Although it is the first of these problems that is suggested by the study of the functional relation, $z = f(x, y)$, we shall need to deal with both of them in the applications of differentiation and integration. We shall consider some simple examples. For brevity, we shall speak of a point satisfying an equation, when what is meant, of course, is that its coordinates satisfy the equation.

EXAMPLE 1. What is the locus of points satisfying the equation $z = 2$?

SOLUTION. Clearly, every point satisfying this equation is 2 units from the xy -plane in the positive direction, i.e., the direction of the positive z -axis. The locus is the plane parallel to the xy -plane and 2 units away, positively.

EXAMPLE 2. What is the locus of points satisfying the equation $x = y$?

SOLUTION. Among the points satisfying this equation are all those in the xy -plane lying on the line bisecting the angle xOy . Let P be such a point (Fig. 6). Then, any point P' , lying on the perpendicular to the xy -plane through P must also satisfy the equation, since the x - and y -coordinates of P' are the same as those of P . Any point not on a perpendicular meeting this line must fail to satisfy the equation, since its

projection on the xy -plane fails to satisfy it. Hence, the locus is the whole plane perpendicular to the xy -plane and cutting it along the bisector of angle xOy .

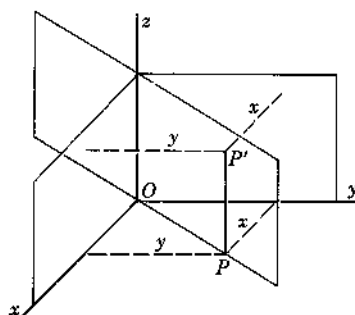


FIG. 6

EXAMPLE 3. What is the equation of the plane perpendicular to the yz -plane and cutting it along the bisector of angle yOz ?

SOLUTION. By comparison with Example 2, we see that the equation of this locus is $y = z$.

EXAMPLE 4. What is the locus of points satisfying the equation, $x^2 + y^2 = 1$?

SOLUTION. Reasoning as in Example 2: (1) we see that the points of this locus that lie in the xy -plane are on the circle of radius 1 with center

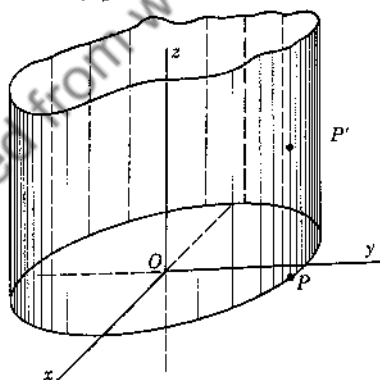


FIG. 7

at 0; (2) if P is one of these points, then any point P' that lies on the line through P and perpendicular to the xy -plane also satisfies the given equation, since its x - and y -coordinates are the same as those of P ; and (3) no point, not on a perpendicular to the xy -plane through a point of the circle, satisfies the equation. Hence, the locus is a circular cylinder of radius 1, perpendicular to the xy -plane (Fig. 7).

EXERCISES

1. What is the equation of the xy -plane? the yz -plane? the xz -plane?
2. Describe the locus of points in space satisfying (a) the equation, $y = 2x$; (b) the equation $x + y = 1$.
3. Describe the locus of points in space satisfying (a) the equation $y^2 + z^2 = 4$; (b) the equation $x^2 + z^2 = 25$.
4. Describe the locus of points in space satisfying (a) the equation $(x - 1)^2 + y^2 = 1$; (b) the equation $y^2 + z^2 - 2y - 2z = 2$.
5. What is the equation of the circular cylinder with elements parallel to the x -axis, having its axis through $(0, 0, 3)$, and its radius 3?
6. Describe the locus of points in space satisfying the equation $(x - 2)^2 + y^2/25 = 1$.
7. A parabolic cylinder has its elements parallel to the z -axis and passing through the curve in the xy -plane: $4y = x^2$. What is its equation?
8. What is the locus in space of points satisfying the equation $yz = 12$?
9. The x -axis is the intersection of two coordinate planes. Its points satisfy what two equations?
10. As in Exercise 9, write two equations satisfied by the y -axis; the z -axis.
11. What is the locus of points satisfying the two equations, $x = 1$ and $y = 2$?
12. What is the locus of points satisfying the two equations, $x^2 + y^2 = 1$ and $z = 3$?

4. Distance from a Coordinate Axis. The directed distances of a point from the coordinate planes are simply the coordinates themselves. How

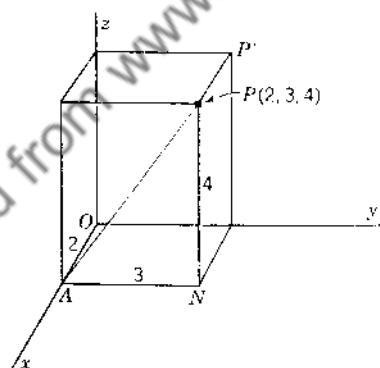


FIG. 8

shall we find the distance of a point from an axis? For example, what is the distance of the point $(2, 3, 4)$ from the x -axis (Fig. 8)? The distance of P from Ox is the length of AP , which is the hypotenuse of right triangle ANP . Since the lengths of AN and NP are 3 and 4, respectively, the length of $AP = \sqrt{3^2 + 4^2} = 5$. In the general case, the point $(2, 3, 4)$ is replaced by (x_1, y_1, z_1) , and the length of $AP = \sqrt{y_1^2 + z_1^2}$. The perpendicular projection of P on the yz -plane, $P'(O, y_1, z_1)$, clearly is the same distance from Ox as is P itself, and so is every point on the line

PP' , since PP' is parallel to Ox . In general, the distances of a point $P(x, y, z)$ from the three axes are, respectively,

$$\left. \begin{aligned} \text{distance from } Ox &= \sqrt{y^2 + z^2}, \\ \text{distance from } Oy &= \sqrt{x^2 + z^2}, \\ \text{distance from } Oz &= \sqrt{x^2 + y^2}. \end{aligned} \right\} \quad (4.1)$$

and

EXERCISES

1. State the distance of each of the following points from each of the coordinate planes, and from each of the coordinate axes: (a) (3, 4, 7); (b) (-5, 12, 0); (c) (9, -12, 16); (d) (0, 0, -11); (e) (-6, -5, -8).

2. Show that the equation of a circular cylinder with axis on the z -axis, and radius a , is $x^2 + y^2 = a^2$.

3. What is the locus of points whose coordinates satisfy the equation $x^2 + z^2 = a^2$? the equation $y^2 + z^2 = a^2$?

5. Distance between Two Points. The distance of a point $P(x, y, z)$ from the origin $O(0, 0, 0)$ is now easily found. We have found that, in Fig. 9, $AP = \sqrt{y_1^2 + z_1^2}$, where AP is perpendicular to Ox . Since the

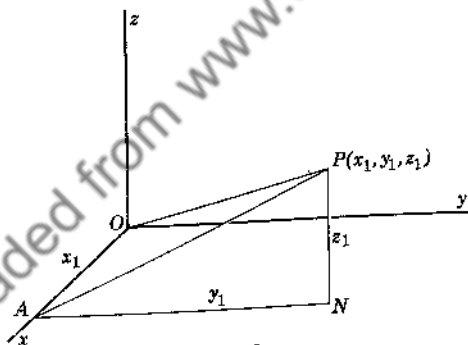


FIG. 9

line OP is the hypotenuse of the right triangle OAP , with $OA = x_1$, we have

$$OP = \sqrt{x_1^2 + y_1^2 + z_1^2}. \quad (5.1)$$

In order to find the distance between two points $Q(x_1, y_1, z_1)$ and $P(x_2, y_2, z_2)$, let us draw Qx' , Qy' , and Qz' through Q , parallel to Ox , Oy , and Oz , respectively (Fig. 10). We thus form a new coordinate system with origin at Q , which is identical with the system obtained by translating the original system to the new origin Q . In this new coordinate system, the coordinates of P may be called (x_2', y_2', z_2') . The distance QP is

therefore equal to $\sqrt{x_2'^2 + y_2'^2 + z_2'^2}$, as above. It is shown in coordinate geometry of two dimensions that if the coordinate system is translated to a new origin (x_1, y_1) , a point $P(x_2, y_2)$ has its new abscissa equal to $(x_2 - x_1)$ and its new ordinate equal to $(y_2 - y_1)$. The demonstration is the same in three dimensions, and need not be repeated. It follows that $x_2' = x_2 - x_1$, $y_2' = y_2 - y_1$, and $z_2' = z_2 - z_1$, and the distance between $Q(x_1, y_1, z_1)$ and $P(x_2, y_2, z_2)$ is

$$\text{distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \quad (5.2)$$

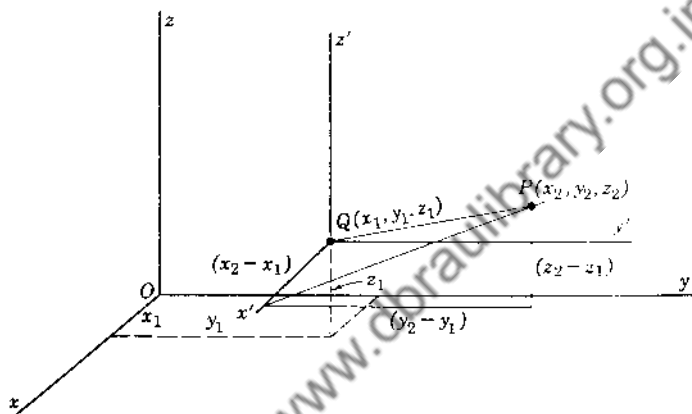


FIG. 10

EXERCISES

- Find the distance of each of the points in Exercise 1, Art. 4, from the origin.
- A tetrahedron is formed by the coordinate planes and the plane through $(0, 0, 9)$, $(0, 12, 0)$, and $(16, 0, 0)$. Find the lengths of the six edges.
- The four vertices of a tetrahedron are $(-1, 3, 10)$, $(0, -3, 4)$, $(4, -7, 1)$, and $(2, 5, -5)$. Find the lengths of the six edges.
- Find the perimeter of the triangle with vertices $(12, 0, 0)$, $(0, -16, 0)$, and $(0, 0, 9)$.
- Find the equation of the locus of points equidistant from $(0, 0, 0)$ and $(5, 5, 5)$.
- Find the equation satisfied by all points 1 unit from the origin.
- Find the equation satisfied by all points 10 units from the origin.
- Which of the following points are less than 10 units from the origin? $(5, 5, 5)$; $(-9, 4, 2)$; $(6, -6, -5)$; $(-1, 7, 7)$; $(3, 3, 9)$.
- Which of the following points are inside the cylinder $x^2 + y^2 = 100$? $(1, 1, 11)$; $(-8, 7, 1)$; $(7, 7, 2)$; $(-3, -4, -10)$.
- Show that the three points $(1, 3, -2)$, $(3, 6, -1)$, and $(-3, -3, -4)$ lie on one line.

6. The Equation of a Sphere. From the results of Art. 5 it follows immediately that the locus of points satisfying the equation

$$x^2 + y^2 + z^2 = a^2 \quad (6.1)$$

consists of all points that are a units from the origin, and, hence, is the surface of the sphere of radius a with center at the origin. Using the word "sphere" in the sense of "spherical surface," we say that (6.1) is the equation of this sphere.

Transposing in (6.1) and extracting square roots, we have

$$z = \pm \sqrt{a^2 - x^2 - y^2}. \quad (6.2)$$

Choosing one part of the double sign $\pm$ to avoid ambiguity, say $+$, we write

$$z = + \sqrt{a^2 - x^2 - y^2}. \quad (6.3)$$

The locus of Eq. 6.2, in which, for each set of values of x and y , z must be given two values, one the negative of the other, is the sphere of Eq. 6.1. The locus of Eq. 6.3, in which z must always have the positive value of the square root, is the hemisphere lying above the xy -plane (Fig. 11).

Equation 6.3 is an example of a functional relation of the form $z = f(x, y)$. The locus of this equation is seen to be a surface (Fig. 11). In general, the locus in three dimensions of $z = f(x, y)$ is a surface, whereas the locus in two dimensions of $y = g(x)$ is a curve.

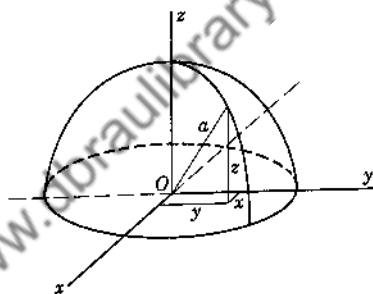


FIG. 11

The sphere with radius a and center at (x_1, y_1, z_1) consists of all the points (x, y, z) at distance a from (x_1, y_1, z_1) . Hence, from (5.2),

$$\sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2} = a,$$

or

$$(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = a^2, \quad (6.4)$$

which is the equation of the given sphere.

EXERCISES

- Write the equation of the sphere with center at $(-2, 1, 3)$ and radius 4, in form (6.4). Multiply out the parenthetical expressions.
- Find the center and radius of the sphere whose equation is $x^2 + y^2 + z^2 - 2x - 10y - 6z - 1 = 0$.
- Find the center and radius of the sphere whose equation is $x^2 + y^2 + z^2 - 4x + 6y + 4z - 47 = 0$.
- A function $f(x, y)$ is represented by the upper half of the sphere in Exercise 2. Write $f(x, y)$.
- A function $z = f(x, y)$ is represented by the lower half of the sphere in Exercise 3. Write $f(x, y)$.

7. Surfaces of Revolution. Many space figures that are important in the applications of calculus in mechanics and engineering are figures of revolution. Several of these were encountered in Chapters 9 and 14. They include circular cylinders, circular cones, and, of course, spheres. In general, a figure of revolution is one that is generated by rotating a plane figure about an axis. For a surface of revolution, the plane figure is a curve. A volume of revolution, on the other hand, is generated by

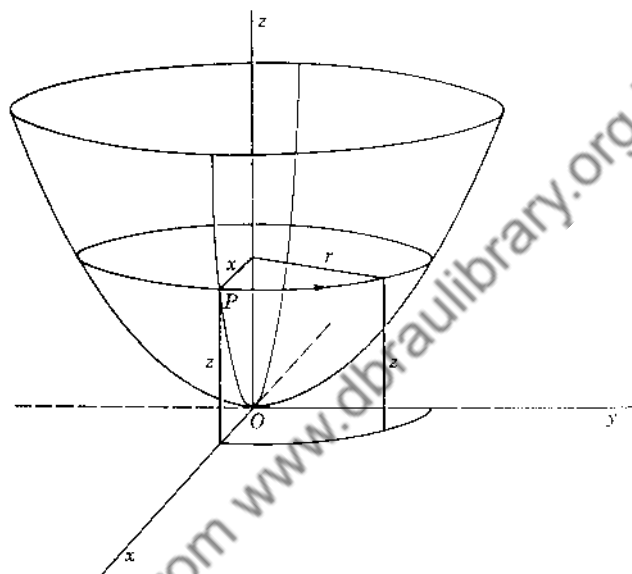


FIG. 12

the rotation of an area. We shall consider only cases in which the axis of the figure is one of the coordinate axes.

Let the parabola in the xz -plane, whose equation is $z = x^2$, be rotated about the z -axis (Fig. 12). Each point of the curve traces a circle, every point of which has the same z -coordinate. Since the radius of this circle is equal to the x -coordinate of P on the parabola, we see that, on the circle, $z = r^2$. Since $r^2 = x^2 + y^2$, by (4.1), we have

$$z = x^2 + y^2. \quad (7.1)$$

The point P is any point on the parabola. Hence, (7.1) is the equation satisfied by all points of the surface, which is called a *paraboloid of revolution*.

By a similar argument, it can be shown that the equation of the surface of revolution generated by rotating about the z -axis the curve (in the xz -plane), $z = f(x)$, is

$$z = f(r) = f(\sqrt{x^2 + y^2}). \quad (7.2)$$

If the curve is $z = g(y)$, in the yz -plane, the equation of the surface is

$$z = g(r) = g(\sqrt{x^2 + y^2}). \quad (7.3)$$

Equations similar in form to (7.2) and (7.3) can be found for surfaces generated by revolving curves about the x -axis or the y -axis (see Exercises 7 and 8, below).

A more general equation of a surface of revolution with Oz as its axis has the form

$$F(z, r) = 0, \quad \text{or} \quad F(z, \sqrt{x^2 + y^2}) = 0. \quad (7.4)$$

The form $F(z, r)$ does not necessarily imply that both variables are present, or that it is possible to solve for one variable in terms of the other. If z is not actually present in $F(z, \sqrt{x^2 + y^2})$, (7.4) becomes the equation of a cylinder (or more than one cylinder), with axis on the z -axis. For example, $x^2 + y^2 - 1 = 0$ (Example 4, Art. 3) has this form. This equation is included in the general form (7.4), but does not have the form $z = f(\sqrt{x^2 + y^2})$.

EXERCISES

1. Obtain the equation of the unit sphere with center at the origin, as a surface of revolution.
2. Obtain the equation of the circular cone generated by revolving the line (in the yz -plane) $z = 3y$ about the z -axis.
3. Write the equation of the paraboloid of revolution generated by revolving the parabola $z = 1 - y^2$ about the z -axis.
4. Describe the locus of points satisfying the equation $z = e^{x^2 + y^2}$.
5. Describe the locus of points satisfying the equation $z^2 = (x^2 + y^2)^2$.
6. Describe the locus of points satisfying the equation $(x^2 + y^2)^2 - 3(x^2 + y^2) + 2 = 0$.
7. Describe the locus of points satisfying the equation $y = x^2 + z^2$.
8. Describe the locus of points satisfying the equation $x = 4 - y^2 - z^2$.
9. Describe the locus of points satisfying the equation $4z^2 - x^2 - y^2 = 0$.
10. An equation of the form $F(z, \sqrt{x^2 + y^2}) = 0$, in which z is not actually present, does not represent z as a function of x and y , since it does not determine a value of z for each pair of values of x and y . How is this brought out by the graphical representation?
11. Obtain the equation of the cone with apex at the origin, axis on the y -axis, and the apex angle 30° . [Hint: It is generated by a line in the xy -plane with slope $= \tan 60^\circ$.]
12. Show that $\sin z = z^2(x^2 + y^2)$ is the equation of a surface of revolution, even though it cannot be solved explicitly for z .

8. Some Special Surfaces. The three-dimensional counterpart of conics in two dimensions are the *quadric surfaces*. These are the surfaces associated with second-degree equations in three variables. We shall illustrate these surfaces, only.

ELLIPSOIDS. Figure 13 shows an ellipsoid centered at the origin, with the equation $x^2/9 + y^2/64 + z^2/16 = 1$. The general form of the equation of an ellipsoid in this position is $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$. If any two of the quantities a , b , and c are equal, the surface is an ellipsoid of revolution. If all three are equal, it is a sphere.

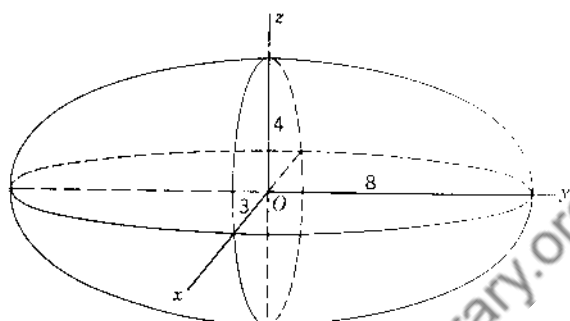


FIG. 13

PARABOLOIDS. There are two paraboloids, an *elliptic paraboloid* has an equation like $z = ax^2 + by^2$, in which a and b have the same sign. This surface looks much like the paraboloid of revolution (Fig. 12), cross sections perpendicular to the z -axis being ellipses. Centered around the y -axis, this surface has an equation of the form $y = ax^2 + bz^2$, and about the x -axis, the equation $x = ay^2 + bz^2$.

Figure 14 shows a *hyperbolic paraboloid*, with the equation $y^2 - x^2 = z$. Planes parallel to the xy -plane intersect this surface in hyperbolas. Those

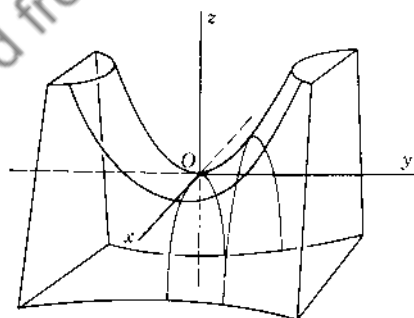


FIG. 14

parallel to the xz -plane or the yz -plane intersect it in parabolas. Another equation for a surface of this kind is $z = axy$, which can be obtained from the general form, $y^2/b^2 - x^2/c^2 = z$, illustrated in Fig. 14, by rotation of the xy -axes. This surface is also called a *saddle surface*, for obvious reasons.

HYPERBOLOIDS. Figures 15 and 16 illustrate hyperboloids of one sheet and two sheets, respectively. The equation of the surface in Fig. 15 has the form $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$. The hyperboloid of two sheets has an equation of the form $z^2/c^2 - x^2/a^2 - y^2/b^2 = 1$.

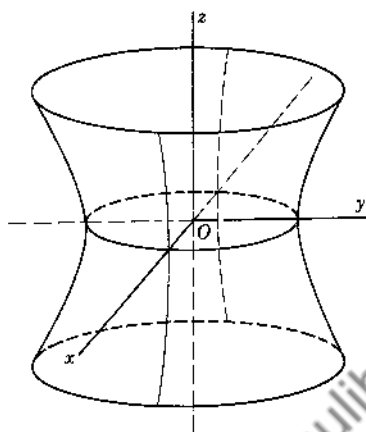


FIG. 15

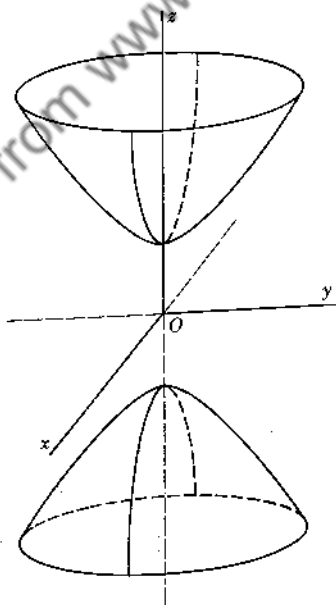


FIG. 16

EXERCISES

1. By analogy with the equation of the elliptic paraboloid discussed above, write the equation of an elliptic paraboloid with vertex at the origin and axis on the x -axis; axis on the y -axis.

2. Sketch the elliptic paraboloid

$$4y + z^2 + 2x^2 = 0.$$

3. Show that the intersection of the hyperbolic paraboloid of Fig. 14 by the xy -plane is two straight lines.

4. Describe the intersection of the hyperboloid of one sheet in Fig. 15, by planes parallel to each of the coordinate planes, respectively.

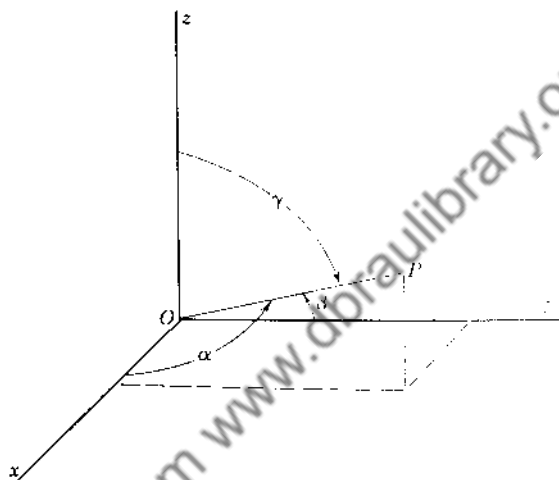


FIG. 17

9. Direction Cosines: Lines through the Origin. In plane coordinate geometry the direction of a line is designated by a function of one angle, viz., the tangent of the angle of inclination of the line to the x -axis. The situation is naturally more complicated in three dimensions. In this case, at least two angles of inclination must be specified, and there are many advantages in the use of three.

In order to narrow the field of view, let us confine our attention, at first, to lines emanating from the origin. Such a line makes three angles with the three coordinate axes, respectively (Fig. 17). Let us denote the angles which the line makes with the x -, y -, and z -axes by α , β , and γ , respectively. These angles are called the *direction angles* of the line OP . For most purposes, it is advantageous to use the cosines of these angles, rather than the angles themselves. (The cosine is more convenient here than the tangent, which plays the dominant role in two dimensions.) The quantities $\cos \alpha$, $\cos \beta$, $\cos \gamma$, are called the *direction cosines* of the

line OP . It should be clear that if the direction cosines are known, the direction is completely determined, even to the sense along the line, i.e., the distinction between the direction OP and PO . For the direction PO , in the figure, the direction angles differ from α, β, γ , respectively, by 180° , and, hence, the direction cosines for the one direction are the negatives of those for the other.

The knowledge of all three direction cosines is a little more than is needed, since there is an identity connecting these three quantities. This relation can be found from Fig. 18, as follows. Let (x_1, y_1, z_1) be the coordinates of point P , and let r = length of OP . Now, AP, BP , and CP

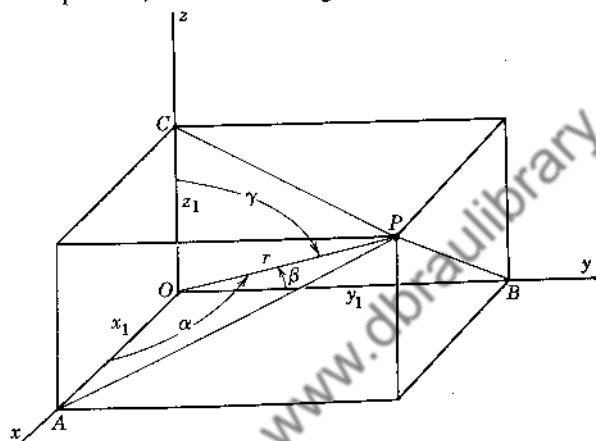


FIG. 18

are perpendicular to Ox, Oy , and Oz , respectively. Hence, OP is the hypotenuse of each of the right triangles OAP, OBP , and OCP . Therefore, $\cos \alpha = x_1/r$, $\cos \beta = y_1/r$, and $\cos \gamma = z_1/r$. Squaring and adding,

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = (x_1^2 + y_1^2 + z_1^2)/r^2.$$

Since

$$r = \sqrt{x_1^2 + y_1^2 + z_1^2},$$

we have

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1. \quad (9.1)$$

Suppose that two lines, OP_1 and OP_2 , emanate from the origin (Fig. 19), the coordinates of P_1 and P_2 being (x_1, y_1, z_1) and (x_2, y_2, z_2) , respectively. Denoting the lengths of OP_1 and OP_2 by r_1 and r_2 , respectively, and the direction angles of OP_1 by $\alpha_1, \beta_1, \gamma_1$, and those of OP_2 by $\alpha_2, \beta_2, \gamma_2$, we have

$$\cos \alpha_1 = x_1/r_1, \quad \cos \beta_1 = y_1/r_1, \quad \cos \gamma_1 = z_1/r_1,$$

$$\cos \alpha_2 = x_2/r_2, \quad \cos \beta_2 = y_2/r_2, \quad \cos \gamma_2 = z_2/r_2.$$

In order to find the angle ω between OP_1 and OP_2 , we apply the law of cosines to triangle OP_1P_2 :

$$\begin{aligned}\cos \omega &= \frac{r_1^2 + r_2^2 - (P_1P_2)^2}{2r_1r_2} \\&= \frac{[x_1^2 + y_1^2 + z_1^2] + [x_2^2 + y_2^2 + z_2^2] - [(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]}{2r_1r_2} \\&= \frac{2(x_1x_2 + y_1y_2 + z_1z_2)}{2r_1r_2} \\&= \frac{x_1}{r_1} \cdot \frac{x_2}{r_2} + \frac{y_1}{r_1} \cdot \frac{y_2}{r_2} + \frac{z_1}{r_1} \cdot \frac{z_2}{r_2}.\end{aligned}$$

Therefore

$$\cos \omega = \cos \alpha_1 \cos \alpha_2 + \cos \beta_1 \cos \beta_2 + \cos \gamma_1 \cos \gamma_2. \quad (9.2)$$

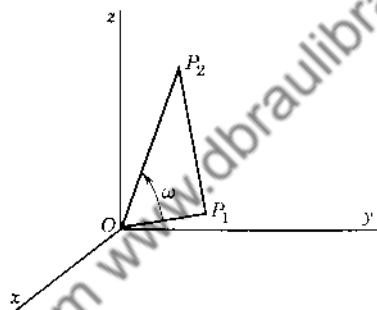


FIG. 19

If $\omega = 90^\circ$, we have

$$\cos \alpha_1 \cos \alpha_2 + \cos \beta_1 \cos \beta_2 + \cos \gamma_1 \cos \gamma_2 = \cos 90^\circ = 0, \quad (9.3)$$

as the condition that OP_1 and OP_2 be perpendicular.

If $\omega = 0$, the lines are coincident and have the same direction; hence

$$\cos \alpha_1 = \cos \alpha_2, \quad \cos \beta_1 = \cos \beta_2, \quad \cos \gamma_1 = \cos \gamma_2. \quad (9.4)$$

If $\omega = 180^\circ$, the lines are coincident but are oppositely directed. Here,

$$\cos \alpha_1 = -\cos \alpha_2, \quad \cos \beta_1 = -\cos \beta_2, \quad \cos \gamma_1 = -\cos \gamma_2. \quad (9.5)$$

Equations 9.4 and 9.5 can also be obtained from 9.1 and 9.2. These derivations are left as exercises for the student (Exercises 5 and 6).

It is sometimes convenient to use, in place of direction cosines, a set of numbers proportional to them. Thus, for example, the perpendicularity condition (9.3) is unchanged by use of such *direction numbers* instead of

direction cosines. If $a_1 : b_1 : c_1 = \cos \alpha_1 : \cos \beta_1 : \cos \gamma_1$, and $a_2 : b_2 : c_2 = \cos \alpha_2 : \cos \beta_2 : \cos \gamma_2$, then

$$a_1 \cdot a_2 + b_1 \cdot b_2 + c_1 \cdot c_2 = 0 \quad (9.6)$$

if, and only if, (9.3) is satisfied.

Direction numbers are sometimes more readily obtained than direction cosines. For example the coordinates of any point on a line through the origin (except the origin itself) are direction numbers of the line. From any set of direction numbers the direction cosines can always be found. Thus, if a, b, c are direction numbers, then $a/r, b/r, c/r$ are the direction cosines, where $r = \sqrt{a^2 + b^2 + c^2}$. The student can prove this (Exercise 2).

EXAMPLE. Find the angle between OP_1 and OP_2 if $(2, 3, 1)$ and $(1, -1, 5)$ are the coordinates of P_1 and P_2 , respectively.

SOLUTION. We have 2, 3, 1 and 1, -1, 5 as direction numbers of OP_1 and OP_2 , respectively. Hence, $2/\sqrt{14}, 3/\sqrt{14}, 1/\sqrt{14}$ and $1/\sqrt{27}, -1/\sqrt{27}, 5/\sqrt{27}$ are the direction cosines, respectively, of these two lines. Therefore, by (9.2),

$$\cos \omega = \frac{2}{\sqrt{14}} \frac{1}{\sqrt{27}} + \frac{3}{\sqrt{14}} \frac{-1}{\sqrt{27}} + \frac{1}{\sqrt{14}} \frac{5}{\sqrt{27}} = \frac{4}{3\sqrt{42}}$$

and

$$\omega = 78^\circ 7' 39''.$$

EXERCISES

- Find the direction cosines of OP for each of the following locations of P : (a) $(1, 1, 1)$; (b) $(-1, -1, -1)$; (c) $(2, 3, 0)$; (d) $(-1, \sqrt{3}, -\sqrt{5})$; (e) $(0, 0, 2)$.
- Show that if a, b, c are direction numbers, then $a/r, b/r, c/r$ are the direction cosines, where $r = \sqrt{a^2 + b^2 + c^2}$.

3. Show that OP_1 and OP_2 are perpendicular if, and only if, $x_1x_2 + y_1y_2 + z_1z_2 = 0$, where the coordinates of P_1 are (x_1, y_1, z_1) and those of P_2 are (x_2, y_2, z_2) .

- Test OP_1 and OP_2 for perpendicularity.

(a) $P_1(3, 0, -2), P_2(-6, 9, -9)$.

(b) $P_1(0, 0, 3), P_2(8, -1, 0)$.

(c) $P_1(a, 0, c), P_2(0, p, 0)$.

(d) $P_1(5, -2, 7), P_2(-3, -7, 1)$.

(e) $P_1(-3, -5, 6), P_2(5, 3, 5)$.

5. Derive Eqs. 9.4 from the relations $\cos^2 \alpha_1 + \cos^2 \beta_1 + \cos^2 \gamma_1 = 1$; $\cos^2 \alpha_2 + \cos^2 \beta_2 + \cos^2 \gamma_2 = 1$; and $\cos \alpha_1 \cos \alpha_2 + \cos \beta_1 \cos \beta_2 + \cos \gamma_1 \cos \gamma_2 = 1$. [Hint: Add the first two equations and - 2 times the third.]

6. Derive Eqs. 9.5 from the relations $\cos^2 \alpha_1 + \cos^2 \beta_1 + \cos^2 \gamma_1 = 1$; $\cos^2 \alpha_2 + \cos^2 \beta_2 + \cos^2 \gamma_2 = 1$; and $\cos \alpha_1 \cos \alpha_2 + \cos \beta_1 \cos \beta_2 + \cos \gamma_1 \cos \gamma_2 = -1$.

7. Write the direction cosines of the coordinate axes.
8. Find the angle between OP_1 and OP_2 if $(0, 4, 3)$ and $(-2, 2, 5)$ are the coordinates of P_1 and P_2 , respectively.
9. How can three angles be tested as possible direction angles of a line? By this test, find which of the following sets of angles are possible direction angles. (a) $30^\circ, 60^\circ, 90^\circ$. (b) $0^\circ, 45^\circ, 45^\circ$. (c) $90^\circ, 45^\circ, 45^\circ$. (d) $120^\circ, 150^\circ, 90^\circ$. (e) $120^\circ, 60^\circ, 45^\circ$. (f) $45^\circ, 0^\circ, 90^\circ$.
10. Two lines have direction numbers 1, 1, 1 and 4, -3, 0, respectively. Find direction numbers of a line perpendicular to both.

10. Direction Cosines: Any Lines. If a line does not pass through the origin, we imagine a translation of coordinates, setting up new axes through some point of the line. The direction angles, α, β, γ of the line

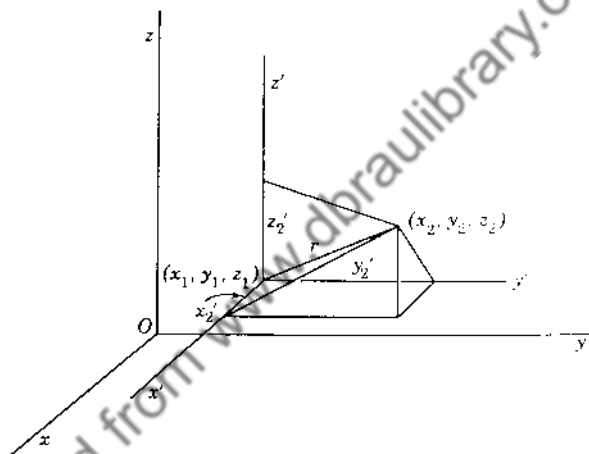


FIG. 20

are the angles that the line makes with the new x -, y -, and z -axes, respectively. The cosine of the angle between two intersecting lines is again found by formula 9.2. If two lines do not intersect, we still speak of the angle between them, meaning the angle at which they would intersect if one were moved parallel to itself so as to meet the other. Formula 9.2 still gives the cosine of this angle.

Two lines are said to be perpendicular if their direction cosines satisfy Eq. 9.3, whether or not they meet. That is, they are called perpendicular if it is possible to translate one of them so that it intersects the other at right angles. Thus, for example, any line in the xy -plane is perpendicular to the z -axis. Two lines are parallel if their direction cosines satisfy either Eqs. 9.4 or Eqs. 9.5, or if direction numbers of one are proportional to direction numbers of the other.

Direction numbers of any line can be easily found from the coordinates of any two points on it. Thus, if a line passes through (x_1, y_1, z_1) and (x_2, y_2, z_2) , we imagine the axes as translated to one of these points, say (x_1, y_1, z_1) , making the new coordinates of the second point $x_2' = x_2 - x_1$, $y_2' = y_2 - y_1$, and $z_2' = z_2 - z_1$ (Fig. 20). Now, just as x_1, y_1, z_1 are direction numbers of a line through the origin and the point (x_1, y_1, z_1) , so x_2', y_2', z_2' are direction numbers of the line in question. That is, it has the direction numbers $(x_2 - x_1), (y_2 - y_1), (z_2 - z_1)$. Its direction cosines are, therefore,

$$\frac{x_2 - x_1}{r}, \frac{y_2 - y_1}{r}, \frac{z_2 - z_1}{r},$$

where

$$r = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

EXERCISES

1. Write the direction cosines of lines parallel to the coordinate axes.
2. What is $\cos \gamma$ for a line in the xy -plane? What can be said about a line in the xz -plane? in the yz -plane?
3. Show that $(3, 2, 5)$, $(6, 5, 4)$, and $(5, 3, -5)$ are vertices of a right triangle.
4. What kind of quadrilateral has the vertices $(1, 2, 3)$, $(5, 3, 6)$, $(6, 4, 11)$, and $(2, 3, 8)$?
5. Which of the following sets of points lie on a straight line?
 - (a) $(2, 3, 4)$, $(6, 5, 8)$, $(10, 7, 12)$.
 - (b) $(2, -3, 4)$, $(6, 8, -2)$, $(-3, -4, 5)$.
 - (c) $(3, 4, -2)$, $(5, -3, 6)$, $(7, -10, 14)$.
 - (d) $(2, 0, -5)$, $(-3, 0, 6)$, $(4, 0, 2)$.

11. Planes and their Equations. In Art. 3 we found the equations of planes parallel to the coordinate planes, and those that contain a coordinate axis. We can make the following general statements.

1. The plane parallel to the xy -plane and intersecting the z -axis at $z = k$ has the equation $z = k$. Similar statements can be made concerning planes parallel to the xz -plane and the yz -plane.
2. A plane containing the z -axis and meeting the xy -plane in the line $ax + by = 0$ has the same equation, viz., $ax + by = 0$. We note that when the space under consideration is limited to the xy -plane, the equation $ax + by = 0$ is the equation of a line in this plane; but when our view encompasses the space of three dimensions, this same equation is the equation of the plane whose projection on the xy -plane is the above line. Similar statements can be made concerning planes containing the x -axis or the y -axis.

We now consider planes more generally. Just as in two dimensions a line is determined by one of its points and its slope, or the slope of its normal, so in three dimensions a plane is determined by one of its points and the direction normal to the plane.

EXAMPLE 1. Suppose a plane passes through $(2, 3, 5)$ and that direction numbers of a normal are $1, -5, 4$. Let (x, y, z) be any point of the plane other than $(2, 3, 5)$ (Fig. 21). Then, $(x - 2), (y - 3), (z - 5)$ are direction numbers of the line, in the plane, that joins $(2, 3, 5)$ and (x, y, z) .

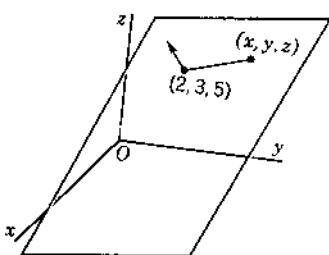


FIG. 21

This line is perpendicular to the given normal direction, since the normal is perpendicular to every line in the plane. Hence, by (9.6),

$$1 \cdot (x - 2) + (-5) \cdot (y - 3) + 4 \cdot (z - 5) = 0. \quad (11.1)$$

Since (11.1) is obviously satisfied by $(2, 3, 5)$, it is satisfied by every point in the plane.

Conversely, if (x, y, z) is any point satisfying (11.1), the line joining it to $(2, 3, 5)$ is perpendicular to the given normal direction, and, hence, lies in the plane. That is, (11.1) is satisfied by all points in the plane and by no others.

In precisely the same way, the plane through (x_1, y_1, z_1) , with normal direction determined by the direction numbers (a, b, c) , has the equation

$$a(x - x_1) + b(y - y_1) + c(z - z_1) = 0. \quad (11.2)$$

Since every plane has a definite normal direction and any one of its points can be taken as (x_1, y_1, z_1) , it follows that every plane has an equation of the form (11.2). In particular, the equation of a plane is of the first degree in x, y , and z .

Conversely, every equation of the first degree in x, y , and z has a plane for its locus. The general equation of the first degree is

$$Ax + By + Cz + D = 0. \quad (11.3)$$

We can find three values x_1, y_1, z_1 , satisfying this equation, in many ways, e.g., by taking x_1 and y_1 arbitrarily and solving (11.3) for the corresponding value of z_1 . We then have

$$Ax_1 + By_1 + Cz_1 + D = 0. \quad (11.4)$$

Subtracting (11.4) from (11.3),

$$A(x - x_1) + B(y - y_1) + C(z - z_1) = 0. \quad (11.5)$$

Equation 11.5 has the form of (11.2) and therefore is the equation of the plane through (x_1, y_1, z_1) with A, B, C as direction numbers of its normal.

Since (11.3) and (11.5) are satisfied by the same values, (11.3) is also the equation of the plane.

It is of special interest that in the general plane equation (11.3), the coefficients A , B , C are direction numbers of the normal. The direction cosines of the normal are therefore

$$\frac{A}{\pm \sqrt{A^2 + B^2 + C^2}}, \quad \frac{B}{\pm \sqrt{A^2 + B^2 + C^2}}, \quad \frac{C}{\pm \sqrt{A^2 + B^2 + C^2}},$$

the choice of sign determining the positive sense on the normal.

The equation of the plane determined by three points, not all on the same straight line, can be obtained from (11.3). This is illustrated in Example 2.

EXAMPLE 2. Find the equation of the plane through $(2, 1, 0)$, $(-4, 3, 1)$, and $(3, -5, -2)$.

SOLUTION. Let the required equation be

$$Ax + By + Cz + D = 0, \quad (11.6)$$

where A , B , C , and D are to be determined. We think of A , B , C , and D as the unknowns, and write (11.6) in the form

$$xA + yB + zC + D = 0. \quad (11.7)$$

Since (11.7) must be satisfied by each of the given points, we then have

$$2A + 1B + 0C + 1D = 0, \quad (11.8)$$

$$-4A + 3B - 1C + 1D = 0, \quad (11.9)$$

and

$$3A - 5B - 2C + 1D = 0. \quad (11.10)$$

We know that there is a plane through the given points, since a plane can be passed through any three points. Hence, there must exist numerical values of A , B , C , and D , not all 0, satisfying (11.8), (11.9), (11.10), and also (11.7), where (x, y, z) is any point of the plane. From algebra, the condition for the existence of such a solution is

$$\begin{vmatrix} x & y & z & 1 \\ 2 & 1 & 0 & 1 \\ -4 & 3 & -1 & 1 \\ 3 & -5 & -2 & 1 \end{vmatrix} = 0. \quad (11.11)$$

This is the required equation. From the definition of a determinant we know that the equation is of the first degree, and therefore represents a

plane. As a check of the fact that this plane passes through the given points, we substitute the coordinates of each point, one after the other, for x , y , and z in the determinant. In each case the determinant becomes 0 because of the presence of identical rows. Expanding the determinant and changing signs, Eq. 11.11 becomes

$$10x + 13y - 34z - 33 = 0.$$

EXERCISES

1. Find direction cosines of a normal to the plane $5x + y + 2z + 3 = 0$.
2. Find the intercepts of the plane $2x + y - 3z - 12 = 0$ on the coordinate axes, and use them to draw the plane.
3. Find the equation of the plane through (2, 2, 2) that cuts off equal intercepts on the coordinate axes.
4. Write the equation of the plane through (2, 1, 1), (1, -1, 0), and (5, 3, -4).
5. Show that the line through (-1, 8, 1) with direction numbers 5/8, -5, -1 lies in the plane $8x + 2y - 5z - 3 = 0$.
6. Show that the planes $ax + by + cz + d = 0$ and $Ax + By + Cz + D = 0$ are parallel if, and only if, $a/A = b/B = c/C$. What condition must be added to make the planes coincident?
7. Show that the planes $3x - 2y + 13z - 15 = 0$ and $7x + 4y + z + 11 = 0$ are perpendicular.
8. Let a normal to a plane be drawn through the origin, meeting the plane at $P(x_1, y_1, z_1)$, and let $p =$ the length of OP . Let (x, y, z) be any other point in the plane. Show that

$$(x - x_1)x_1 + (y - y_1)y_1 + (z - z_1)z_1 = 0$$
 and, hence,

$$x \cos \alpha + y \cos \beta + z \cos \gamma - p = 0, \quad (11.12)$$
 where α, β, γ are direction angles of the normal and $p = \sqrt{x_1^2 + y_1^2 + z_1^2}$. Equation 11.12 is the normal form of the equation of a plane.
9. If (x_2, y_2, z_2) is a point not on the plane represented by (11.12), show that the distance of this point from that plane is equal to

$$x_2 \cos \alpha + y_2 \cos \beta + z_2 \cos \gamma - p,$$
 with the sign positive if, and only if, the plane lies between the point (x_2, y_2, z_2) and the origin.
10. Use the result of Exercise 9 to find the distance of the point (4, 3, -7) from the plane $10x + 4y - 5z - 20 = 0$.

12. Straight Lines. Since a straight line is determined as the intersection of two planes, every point on it must satisfy two plane equations. This is analogous to the determination of the coordinates of a point in two dimensions as the solution of two simultaneous linear equations. We therefore speak of the equations (plural) of a line. We consider three important forms these equations may take.

GENERAL FORM. As an example, we take the line given by the equations

$$\begin{aligned} 3x - 2y - 4z + 2 &= 0; \\ 2x + 5y - 2z - 7 &= 0. \end{aligned} \quad (12.1)$$

The planes represented by these equations are not parallel (see Exercise 6, Art. 11), and, hence, meet in a line. Suppose l, m, n are direction numbers of this line. Then, since the line is perpendicular to the normals of both planes,

$$\begin{aligned} 3l - 2m - 4n &= 0 \\ 2l + 5m - 2n &= 0. \end{aligned} \quad (12.2)$$

Now, since a determinant with two identical rows is equal to 0,

$$\begin{vmatrix} 3 & -2 & -4 \\ 3 & -2 & -4 \\ 2 & 5 & -2 \end{vmatrix} = \begin{vmatrix} 2 & 5 & -2 \\ 3 & -2 & -4 \\ 2 & 5 & -2 \end{vmatrix} = 0.$$

Upon expanding each of these determinants by minors of the first row, we have

$$3 \begin{vmatrix} -2 & -4 \\ 5 & -2 \end{vmatrix} + (-2) \begin{vmatrix} -4 & 3 \\ -2 & 2 \end{vmatrix} + (-4) \begin{vmatrix} 3 & -2 \\ 2 & 5 \end{vmatrix} = 0$$

$$2 \begin{vmatrix} -2 & -4 \\ 5 & -2 \end{vmatrix} + 5 \begin{vmatrix} -4 & 3 \\ -2 & 2 \end{vmatrix} + (-2) \begin{vmatrix} 3 & -2 \\ 2 & 5 \end{vmatrix} = 0.$$

Comparing these results with (12.2), we see that Eqs. 12.2 are satisfied by

$$l = \begin{vmatrix} -2 & -4 \\ 5 & -2 \end{vmatrix} = 24, \quad m = \begin{vmatrix} -4 & 3 \\ -2 & 2 \end{vmatrix} = 2, \quad n = \begin{vmatrix} 3 & -2 \\ 2 & 5 \end{vmatrix} = 19.$$

These determinants are obtained from the matrix of the coefficients in (12.2):

$$\begin{pmatrix} 3 & -2 & -4 \\ 2 & 5 & -2 \end{pmatrix}.$$

For l , we strike out the first column; for m , we strike out the second column and write the third and first columns; for n , we strike out the third column and write the first and second. The columns appear in cyclic order, in pairs: second third, third first, first second.

SYMMETRIC FORM. Since infinitely many planes intersect along a given line, the equations of the line are those of any pair of these planes. We can choose this pair, according to convenience, among all planes intersecting in the line. Suppose we wish to find equations of the line through $(2, 1, 4)$ with direction numbers 3, 5, 1. If (x, y, z) is any other point of the line, $x - 2$, $y - 1$, $z - 4$ are also direction numbers of the line. Hence $\frac{x - 2}{3} = \frac{y - 1}{5} = \frac{z - 4}{1}$. This continued equality is equivalent to the two equations

$$\frac{x - 2}{3} = \frac{y - 1}{5} \quad \text{and} \quad \frac{x - 2}{3} = \frac{z - 4}{1} \quad (12.3)$$

which are the equations of two planes intersecting along the given line. The first of these planes is perpendicular to the xy -plane and the second to the xz -plane. The equation $(y - 1)/5 = (z - 4)/1$ can be obtained by combining Eqs. 12.3, and, hence, is not independent of them. Geometrically, this corresponds to the fact that this plane, perpendicular to the yz -plane, intersects both planes (12.3) in the same line.

In general, the line through the point (x_1, y_1, z_1) , with direction numbers a, b, c , has equations

$$\frac{x - x_1}{a} = \frac{y - y_1}{b} = \frac{z - z_1}{c} \quad (12.4)$$

Equations 12.4 are called the *symmetric equations* of the line. Writing the two independent equations

$$\frac{x - x_1}{a} = \frac{y - y_1}{b} \quad \text{and} \quad \frac{x - x_1}{a} = \frac{z - z_1}{c}$$

we obtain, by clearing fractions,

$$b(x - x_1) = a(y - y_1) \quad \text{and} \quad c(x - x_1) = a(z - z_1). \quad (12.5)$$

Even if one of the direction numbers is 0, form (12.4) is still used. Thus, if the direction numbers are 2, 3, 0, and the line passes through $(3, 1, 4)$, we write

$$\frac{x - 3}{2} = \frac{y - 1}{3} = \frac{z - 4}{0}.$$

This does not mean that we have introduced division by 0, but merely that we agree that this form shall symbolize two equations of the form (12.5), with $c = 0$: $3(x - 3) = 2(y - 1)$ and $0(x - 3) = 2(z - 4)$, which reduce to $3x - 2y - 7 = 0$ and $z - 4 = 0$.

PARAMETRIC FORM. Again, let us take the line through (2, 1, 4) with direction numbers 3, 5, 1. We found that the condition that a point (x, y, z) be on the line was the equality of the three ratios in (12.3). Of course, as the point (x, y, z) takes different positions along the line, the value of each ratio changes, while the three remain equal to one another. We therefore write

$$\frac{x-2}{3} = \frac{y-1}{5} = \frac{z-4}{1} = t,$$

where t is the varying value of these equal ratios. From this, we have

$$x-2=3t, \quad y-1=5t, \quad z-4=t$$

or

$$x=2+3t, \quad y=1+5t, \quad z=4+t. \quad (12.6)$$

Equations 12.6 are thus *parametric equations* of the straight line. By eliminating the parameter t among these equations, we would return to two equations like (12.3).

In general, the line through (x_1, y_1, z_1) , with direction numbers a, b, c , has the parametric equations

$$x=x_1+at, \quad y=y_1+bt, \quad z=z_1+ct. \quad (12.7)$$

EXAMPLE 1. Show that (2, -3, -5) lies on the line

$$8x-4y+3z-13=0,$$

$$x+3y-z+2=0.$$

SOLUTION. We show that (2, -3, -5) satisfies the equations of both intersecting planes.

$$8 \cdot 2 - 4(-3) + 3(-5) - 13 = 16 + 12 - 15 - 13 = 0,$$

$$2 + 3(-3) - (-5) + 2 = 2 - 9 + 5 + 2 = 0.$$

EXAMPLE 2. If a line has direction numbers -2, 3, 2, and passes through (5, 0, -4), write its equations in symmetric form and in parametric form.

SOLUTION. From (12.4) we write the symmetric equations,

$$\frac{x-5}{-2} = \frac{y}{3} = \frac{z+4}{2}.$$

Using (12.7), the parametric form is

$$x=5-2t, \quad y=3t, \quad z=-4+2t.$$

EXAMPLE 3. If a line is given parametrically by the equations, $x = 3t$, $y = 1 + t$, $z = 3 - 2t$, find symmetric equations for the line. Hence, draw two planes intersecting in the line.

SOLUTION. From these equations, we have

$$t = x/3, \quad t = y - 1, \quad t = (z - 3)/(-2).$$

Hence, $\frac{x}{3} = \frac{y - 1}{1} = \frac{z - 3}{-2}$. From the first equality we have $x - 3y = -3$, which is the equation of plane P in Fig. 22. From

$$\frac{x}{3} = \frac{z - 3}{-2}, \quad \text{we have } 2x + 3z = 9,$$

which is the equation of plane Q . The line of intersection is MN .

EXAMPLE 4. Find the equations of the line through $(-2, -1, 4)$ and $(3, 1, 7)$.

SOLUTION. The differences, $3 - (-2) = 5$, $1 - (-1) = 2$, and $7 - 4 = 3$, are direction numbers of the given line. Using the symmetric form, with the point $(-2, -1, 4)$, we obtain

$$\frac{x + 2}{5} = \frac{y + 1}{2} = \frac{z - 4}{3}.$$

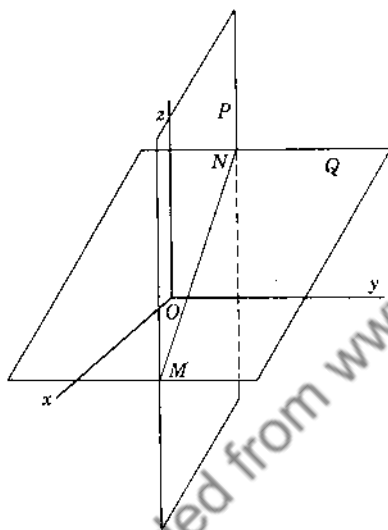


FIG. 22

EXAMPLE 5. Find the plane containing two given intersecting lines.

SOLUTION. We assume the knowledge of direction numbers of both lines, and their point of intersection. In whatever way the lines are given, this knowledge is readily obtained. Suppose the lines meet at (x_1, y_1, z_1) and have direction numbers l_1, m_1, n_1 , and l_2, m_2, n_2 , respectively. Since the plane is to contain both lines, its normal is perpendicular to them both. Hence, if we denote direction numbers of the normal by a, b, c , we have

$$al_1 + bm_1 + cn_1 = 0 \quad (12.8)$$

$$al_2 + bm_2 + cn_2 = 0.$$

Comparing these equations with (12.2), we see that the problem of determining a, b, c is identical in nature with the problem of determining

l , m , and n at that point. Hence, we conclude that Eqs. 12.8 are satisfied by

$$a = \begin{vmatrix} m_1 & n_1 \\ m_2 & n_2 \end{vmatrix}, \quad b = \begin{vmatrix} n_1 & l_1 \\ n_2 & l_2 \end{vmatrix}, \quad c = \begin{vmatrix} l_1 & m_1 \\ l_2 & m_2 \end{vmatrix}.$$

Putting these values in Eq. 11.2, we have the required equation.

EXERCISES

Find equations of each of the lines in Exercises 1-7.

1. Through the origin, intersecting x -, y -, and z -axes at 60° , 120° , and 45° , respectively.
2. Through $(2, -1, 3)$, $\alpha = 45^\circ$, $\beta = 60^\circ$, $\gamma = 120^\circ$.
3. Through $(5, 2, -1)$ and $(2, -3, 4)$.
4. Through $(7, -1, 2)$, perpendicular to the plane $5x + 3y - 6z - 10 = 0$.
5. Through $(3, 0, 1)$, parallel to the line in Exercise 3.
6. Through $(8, 4, -5)$, parallel to the line with the equations $3x - 2y + 1 = 0$, $x + 3y - z = 0$.
7. Lying in the plane $7x - 10y - 4z + 7 = 0$, and meeting the line, $x - y + z = 1$, $4x + 5y - 2z = 7$, perpendicularly.
8. Show that the lines $2x + y - 2z = 12$, $x - 3y + z + 8 = 0$, and $x + 2y + 2z + 5 = 0$, $x + z = 7$, are perpendicular.
9. Find the equations of the line through $(-3, 4, -5)$ and perpendicular to both lines in Exercise 8.
10. Find the equation of the plane containing the lines

$$x = 3 + 2t, y = -1 + 5t, z = 2 - 3t,$$

and

$$x = 3 - t, y = -1 + 2t, z = 2 + 4t.$$

13. Curves in Space. Just as straight lines in space must be looked upon as the intersections of planes, so curves are the intersections of surfaces. Hence, in general, a curve has two equations in the three variables x, y, z . A set of three equations expressing x, y , and z in terms of a fourth variable, or parameter, is equivalent to a pair of equations in x, y , and z , and is often more convenient. Thus, the parametric equations of a curve in space have the form

$$x = \phi(t), \quad y = \psi(t), \quad z = \chi(t).$$

Parametric equations like these, for a curve in a plane, were treated in Chapters 6 and 14.

Some curves that assume a good deal of importance in Chapter 24 are those that are formed by the intersection of a curved surface and a plane. This is illustrated in Example 1.

EXAMPLE 1. The intersection of the elliptic paraboloid $z = x^2/4 + y^2/9$ by the plane $y = 3$ is a parabola, PQR in Fig. 23. If we substitute $y = 3$ in the paraboloid equation, we obtain $z = 1 + x^2/4$, which is the equation of the parabola in its plane.

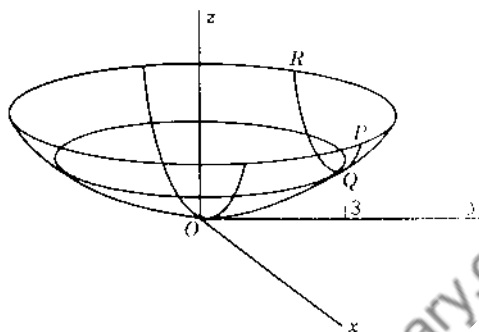


FIG. 23

EXAMPLE 2. The two cylinders, $x^2 + z^2 = 4$ and $y^2 + z^2 = 4$, intersect as shown in Fig. 24. (The figure shows the half of the solid common to the cylinders above the xy -plane, and half of each of the ellipses in which they intersect, AQD and CQB .) The ellipse AQD lies in the plane $x = y$, and CQB in the plane $x = -y$. Hence, the ellipse AQD also

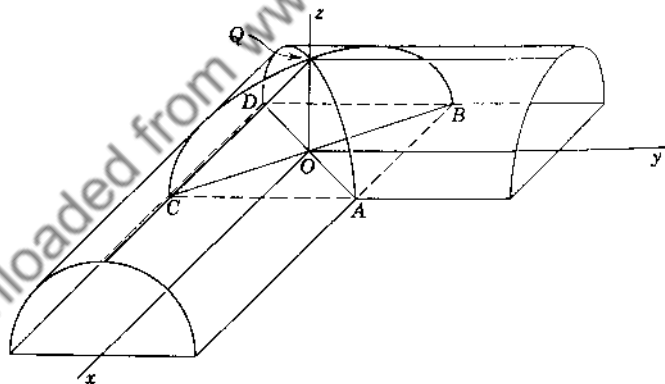


FIG. 24

has the equations $x = y$, $x^2 + z^2 = 4$, and $x = y$, $y^2 + z^2 = 4$. Similarly, the ellipse CQB also has the equations $x = -y$, $x^2 + z^2 = 4$, and $x = -y$, $y^2 + z^2 = 4$.

The parametric equations of a curve in space can be treated in a way entirely analogous to the treatment in a plane (see Chapter 6, Arts. 6 and 7, Chapter 14, Arts. 5, 6, and 7, and Chapter 16, Art. 1). If

$$x = \phi(t), \quad y = \psi(t), \quad z = \chi(t),$$

we can define a vector v , (Fig. 25), with components

$$v_x = dx/dt, \quad v_y = dy/dt, \quad v_z = dz/dt, \quad |v| = \sqrt{v_x^2 + v_y^2 + v_z^2}.$$

(If the variable t represents time, this vector is the velocity of a point along the path.) In addition, we have

$$ds/dt = \pm \sqrt{v_x^2 + v_y^2 + v_z^2} = \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2}, \quad (13.1)$$

$$v_x/v = \frac{dx/dt}{ds/dt} = \cos \alpha,$$

$$v_y/v = \frac{dy/dt}{ds/dt} = \cos \beta,$$

$$v_z/v = \frac{dz/dt}{ds/dt} = \cos \gamma.$$

Hence,

$$dx/dt : dy/dt : dz/dt = \cos \alpha : \cos \beta : \cos \gamma.$$

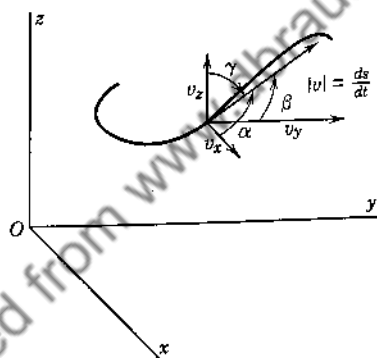


FIG. 25

That is, dx/dt , dy/dt , dz/dt , evaluated at a point of the curve (x_0, y_0, z_0) corresponding to $t = t_0$, viz., $\phi'(t_0)$, $\psi'(t_0)$, $\chi'(t_0)$, are direction numbers of the tangent. It follows from Art. 12 that

$$\frac{x - x_0}{\phi'(t_0)} = \frac{y - y_0}{\psi'(t_0)} = \frac{z - z_0}{\chi'(t_0)} \quad (13.2)$$

are equations of the tangent. Also, the equation of the plane normal to the curve at this point is

$$(x - x_0)\phi'(t_0) + (y - y_0)\psi'(t_0) + (z - z_0)\chi'(t_0) = 0. \quad (13.3)$$

If t is not associated with time, this discussion is still valid, the only difference being that the vector v no longer represents the velocity of a

moving point. It is still the vector with dx/dt , dy/dt , dz/dt as components, and the remainder of the discussion holds since it stems from this idea.

From (13.1) we obtain a formula for the length of arc between two points on the curve corresponding to t_1 and t_2 :

$$s = \left| \int_{t_1}^{t_2} \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt \right|.$$

In this formula we use the absolute value, since length is an unsigned quantity. If it is desired to let the result reflect the sense of description of the curve, being plus in one direction and minus in the other, the absolute value signs may be omitted.

The equations of a space curve of intersection of two cylinders, $y = f(x)$, $z = g(x)$, can be treated as an example of parametric representation, as follows. We let

$$x = t, \quad y = f(t), \quad z = g(t).$$

Equations 13.2 then become

$$(x - x_0)/1 = (y - y_0)/f'(x_0) = (z - z_0)/g'(x_0),$$

and (13.3) becomes

$$(x - x_0) + (y - y_0)f'(x_0) + (z - z_0)g'(x_0) = 0.$$

The length of arc is found by the formula,

$$s = \left| \int_{x_1}^{x_2} \sqrt{1 + (dy/dx)^2 + (dz/dx)^2} dx \right|.$$

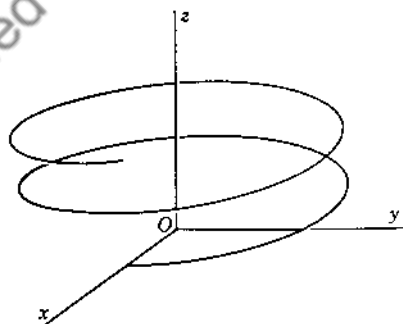


FIG. 26

EXAMPLE 3. The equations, $x = a \cos t$, $y = a \sin t$, $z = bt$ represent a helix, a curve shaped like a coil spring (Fig. 26). Differentiating, we have

$$dx/dt = -a \sin t, \quad dy/dt = a \cos t, \quad dz/dt = b.$$

Let us find equations of the tangent line at $t_0 = \pi/3$. We have

$$x_0 = a \cos(\pi/3) = a/2, \quad y_0 = a \sin(\pi/3) = a\sqrt{3}/2, \quad z_0 = b\pi/3,$$

$$dx/dt = -a \sin(\pi/3) = -a\sqrt{3}/2, \quad dy/dt = a \cos(\pi/3) = a/2, \quad dz/dt = b.$$

Putting these values in (13.2), we have, for the tangent line,

$$\frac{x - a/2}{-a\sqrt{3}/2} = \frac{y - a\sqrt{3}/2}{a/2} = \frac{z - b\pi/3}{b}.$$

From (13.3), the equation of the normal plane is

$$-\frac{a\sqrt{3}}{2} \left(x - \frac{a}{2}\right) + \frac{a}{2} \left(y - \frac{a\sqrt{3}}{2}\right) + b \left(z - b\frac{\pi}{3}\right) = 0.$$

EXERCISES

Describe the curves given by the equations in Exercises 1-6.

1. $z = x^2 + y^2$, $x = 1$.
2. $z = 25 - (x^2 + y^2)$, $y = 3$.
3. $z = 25 - (x^2 + y^2)$, $2x - y = 0$.
4. $x^2 + y^2 + z^2 = 90$, $z = x^2 + y^2$.
5. $x^2 + y^2 + z^2 = 100$, $(x - 5)^2 + y^2 = 25$.
6. $x = y^2 + z^2$, $x^2 + z^2 = 4$.
7. Find equations of the tangent line and the normal plane to the curve, $x = t^2$, $y = t^3$, $z = t^4$, at the point $(1, 1, 1)$.
8. Show that the tangent to the helix in Example 3 always meets the xy -plane at the same angle, and find that angle.
9. Find the length of one coil of the helix of Example 3, i.e., from $t = 0$ to $t = 2\pi$.
10. A coil spring has a diameter of 1 in. and there are 6 turns per in. of length of the coil. Find the length of wire in 4 in. of spring. [Hint: Use the equations of Example 3.]
11. Find equations of the tangent to the curve $x^2 + y^2 = 25$, $z = x^2$, at $(3, 4, 9)$.

14. Other Coordinate Systems. In three dimensions there are two systems of coordinates in common use that are similar to polar coordinates in the plane. One of these, called *cylindrical coordinates*, consists of polar coordinates in one plane, with a rectangular coordinate perpendicular to this plane. For example, we may replace the rectangular coordinates in the xy -plane by polar coordinates, and extend the system to three dimensions by means of the z -coordinate (Fig. 27). The relations connecting this system to rectangular coordinates are

$$x = r \cos \theta$$

$$y = r \sin \theta$$

$$z = z$$

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1} y/x$$

$$z = z.$$

The locus of points for which $r = \text{a constant}$ is a cylinder whose axis lies on the z -axis; that for which $\theta = \text{a constant}$ is a plane containing the z -axis; and that for which $z = \text{a constant}$ is a plane parallel to the $r\theta$ -plane.

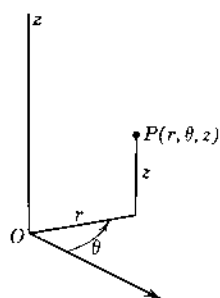


FIG. 27

A second system of space coordinates, bearing a similarity to polar coordinates in the plane, is called *spherical coordinates*, and consists of a radial distance and two angles, as indicated in Fig. 28. In setting up this system we take a point O , and imagine a plane and its directed normal

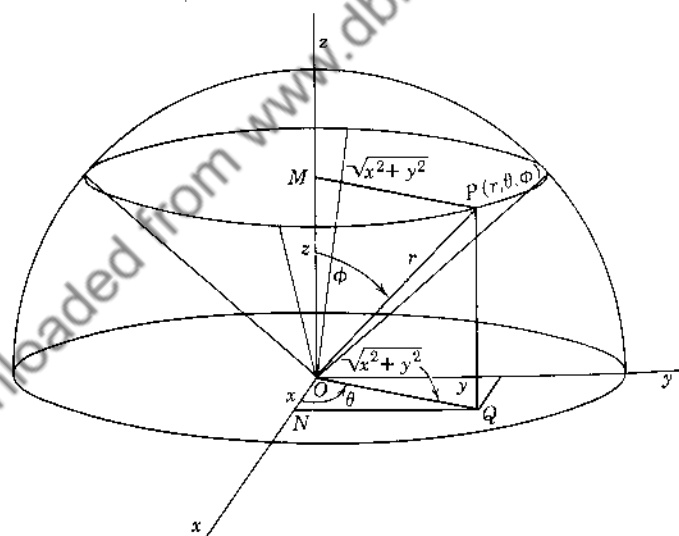


FIG. 28

Oz through O . On this plane we mark off a base direction Ox . Now, for any point P ,

$r = \text{distance } OP$,

$\theta = \text{dihedral angle between plane } xOz \text{ and plane } POz$,

$\phi = \text{angle between } Oz \text{ and } OP$.

The locus of points for which $r = \text{a constant}$ is a sphere; that for which $\theta = \text{a constant}$ is a plane containing Oz ; and that for which $\phi = \text{a constant}$ is a cone with vertex at O and having its axis on Oz .

The relations connecting spherical coordinates with rectangular coordinates are

$$x = r \sin \phi \cos \theta.$$

$$r = \sqrt{x^2 + y^2 + z^2}.$$

$$y = r \sin \phi \sin \theta.$$

$$\theta = \tan^{-1}(y/x).$$

$$z = r \cos \phi.$$

$$\phi = \sin^{-1} \frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}},$$

$$= \cos^{-1} \frac{z}{\sqrt{x^2 + y^2 + z^2}}.$$

These relations can be seen directly from Fig. 28. Thus, for example, to get the relation $x = r \sin \phi \cos \theta$, we note from triangle OMP that

$$\sqrt{x^2 + y^2} = r \sin \phi;$$

from triangle ONQ ,

$$x = \sqrt{x^2 + y^2} \cos \theta = r \sin \phi \cos \theta.$$

Cylindrical and spherical coordinates are most useful in simplifying equations having complicated rectangular forms. Their use in connection with geometric and physical applications in three dimensions, which require integration, can be appreciated only when these problems are reached.

EXERCISES

- Find cylindrical and spherical coordinates of the points whose rectangular coordinates are: $(1, 0, 0)$; $(0, 1, 0)$; $(0, 0, 1)$; $(1, 1, 0)$; $(1, 0, 1)$; $(0, 1, 1)$; $(1, 1, 1)$.
- Find rectangular coordinates of the points whose cylindrical coordinates (r, θ, z) are: $(5, \pi/4, 2)$; $(2, -\pi/2, -3)$; $(4, \pi/3, 5)$.
- Find rectangular coordinates of the points whose spherical coordinates (r, θ, ϕ) are: $(1, \pi/2, \pi/4)$; $(4, 0, 2\pi/3)$; $(2, \pi, 3\pi/4)$; $(10, \pi/6, \pi/2)$.
- Write the rectangular equation of the cone with vertex at the origin, axis on the z -axis, and apex angle $= \pi/3$, by first expressing it in spherical coordinates.
- (a) Transform the equation $x^2 + y^2 + z^2 = 100$ into spherical coordinates. (b) Do the same with the equation $x^2 + y^2 + (z - 1)^2 = 1$.
- Write the equation of the plane $x = k$ in cylindrical coordinates.
- (a) Write the equation of the cylinder $x^2 + y^2 = a^2$ in cylindrical coordinates. (b) Do the same with the cylinder $(x - a)^2 + y^2 = a^2$.
- Write the equation of the sphere $x^2 + y^2 + z^2 = a^2$ in cylindrical coordinates.
- Put the equation of the cone in Exercise 4 into cylindrical coordinates.
- Put the equation $(z^2 + x^2 + y^2)^2 = 4(z^2 - x^2 - y^2)$ into spherical coordinates. Describe the surface of which this is the equation.

15. Summary. It was the intention in this chapter to stress those fundamental ideas of three-dimensional coordinate geometry that have special utility in extending the study of the problems of calculus to functions of more than one variable. In the next chapter we shall pursue the study of the tangent problem, applying it to a function of two variables, where it becomes the problem of determining the tangent plane to a surface. For this purpose an understanding of the equations of planes and of direction cosines has special importance.

Extensions of definite integrals to functions of several variables will be taken up in Chapter 25. There the typical area problem is replaced by a volume problem, a volume under a surface. As in the case of the area problem, numerous applications fall naturally into the framework of this broadened definite integral. For the solution of many of these problems, some knowledge of the equations of surfaces, such as have been briefly discussed in this chapter, is necessary.

Downloaded from www.dbraulibral.com

Functions of Several Variables Differentiation

1. Functions of Two or More Variables. Functions of more than one variable occur frequently in mathematics and its applications. The relation $A = \frac{1}{2}bh$ expresses the area of a triangle as a function of the base and altitude. $A = P(1+i)^n$ expresses the amount of money in an

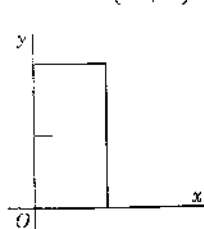


FIG. 1

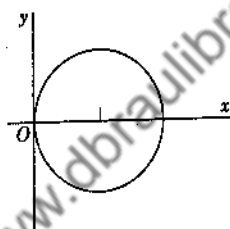


FIG. 2

account as a function of the principal, the interest rate, and the number of interest periods. $E = IR$ expresses the voltage in a circuit as a function of the current and the resistance. In order to deal with practical problems involving such relations and to broaden the scope of calculus in the field of mathematics itself, it is natural to try to extend the powerful operations of differentiation and integration to functions of several variables.

Because of the limitations of space, as well as the restricted scope of this book, we shall treat this material with less detail than was used in presenting the calculus of functions of one variable. The purpose is still, however, to introduce concepts and methods in such a way as to suggest correct formulations, even though the complete development cannot be supplied. Much of the discussion will concern functions of two variables, both because this will make graphical interpretations possible and because results for two variables are readily extended to more than two.

For a function of two variables, $f(x, y)$, the range of the variables is two-dimensional, e.g., the rectangular region: $0 \leq x \leq 1$, $0 \leq y \leq 2$ (Fig. 1); or all points within the circle $(x-1)^2 + y^2 = 1$ (Fig. 2). It may be a region with a more complicated boundary, or even be made up of isolated points. It may be the whole xy -plane, or an infinite part of it,

e.g., all points in the first quadrant. In our work, it will usually be either an infinite region that is easily described, or a region bounded by a single closed curve (which may consist of several arcs, e.g. a square, bounded by four straight lines), or such a region with one point deleted.

We shall consider functions of the sort $z = f(x, y)$, where a single value of z is determined for each pair (x, y) in some region. The notion of a limit of $f(x, y)$ is essentially the same as that of a limit of a function of one variable, although even greater care is needed to determine that a limit actually exists. We say that $f(x, y)$ approaches the limit L as (x, y) approaches (a, b) if we can make $f(x, y)$ differ from L by as little as we please by merely bringing (x, y) sufficiently close to (a, b) , without coinciding with this point. Note that it is only the nearness of (x, y) to (a, b) , regardless of the direction, that guarantees the smallness of $|f(x, y) - L|$. Now, if

1. $f(x, y) \rightarrow L$ as $(x, y) \rightarrow (a, b)$;
2. $f(x, y)$ is defined at (a, b) , i.e., $f(a, b)$ exists;
3. $f(a, b) = L$; then $f(x, y)$ is continuous at (a, b) .

EXAMPLE 1. Let

$$f(x, y) = \frac{x^2 y^2}{x^2 + y^2} \quad \text{if } x^2 + y^2 \neq 0, \\ = 0 \quad \text{if } x^2 + y^2 = 0.$$

We can show that this function is continuous at $(0, 0)$. First, we note that $f(0, 0) = 0$. To examine $\lim f(x, y)$ as $(x, y) \rightarrow (0, 0)$, we use polar coordinates: $x = r \cos \theta$, $y = r \sin \theta$. Hence, if $r \neq 0$, $f(x, y)$ becomes

$$\frac{r^4 \sin^2 \theta \cos^2 \theta}{r^2} = r^2 \sin^2 \theta \cos^2 \theta < r^2.$$

As (x, y) approaches $(0, 0)$, r tends to 0, and, hence, $\lim f(x, y) = 0$.

EXAMPLE 2. Let $f(x, y) = (x - y)/(x^2 + y^2)$, with $x^2 + y^2 \neq 0$. It may be thought that $f(x, y) \rightarrow 0$ as $(x, y) \rightarrow (0, 0)$, since we can always find values of (x, y) making $f(x, y)$ as small as we please, actually 0, by taking $x = y \neq 0$. A special choice of values like this, however, is not enough to prove that $f(x, y) \rightarrow 0$. We would have to show that any choice sufficiently close to $(0, 0)$ makes $f(x, y)$ very small. In fact, this function may approach ∞ as (x, y) tends to $(0, 0)$. Again using polar coordinates,

$f(x, y)$ becomes $\frac{r(\cos \theta - \sin \theta)}{r^2} = \frac{\cos \theta - \sin \theta}{r}$. Now, as $(x, y) \rightarrow (0, 0)$,

the denominator, $r \rightarrow 0$, while the numerator need not tend to 0. Hence, if θ is any value such that $\cos \theta \neq \sin \theta$, $f(x, y) \rightarrow \infty$.

EXAMPLE 3. Let $f(x, y) = xy/(x^2 + y^2)$, with $x^2 + y^2 \neq 0$. This function tends to no single limit L as (x, y) tends to $(0, 0)$. If we let $(x, y) \rightarrow (0, 0)$ through points of the x -axis, the limit is 0; if $(x, y) \rightarrow (0, 0)$ through points on the line $x = y$, the limit is $\frac{1}{2}$, since for these points $xy/(x^2 + y^2) = x^2/(2x^2) = \frac{1}{2}$; if $(x, y) \rightarrow (0, 0)$ through points on the line $y = 2x$, the limit is $\frac{2}{5}$, since for these points $xy/(x^2 + y^2) = 2x^2/(5x^2) = \frac{2}{5}$; etc.

Two different tangent problems suggest themselves in connection with a surface, $z = f(x, y)$. First, there is the question of a tangent to a curve on the surface, and second, the problem of finding the tangent plane to the surface at a given point (x_0, y_0, z_0) . Like the tangent line to a curve $y = f(x)$ at a point (x_0, y_0) , when this problem was first introduced, an important part of the problem is the formulation of a suitable definition of a tangent plane. The first of these problems will be taken up in Art. 2, for the important special cases of curves of intersection of $z = f(x, y)$ by planes $x = x_0$ and $y = y_0$. We shall find the directions of such curves, from which the equations of the tangents can easily be found.

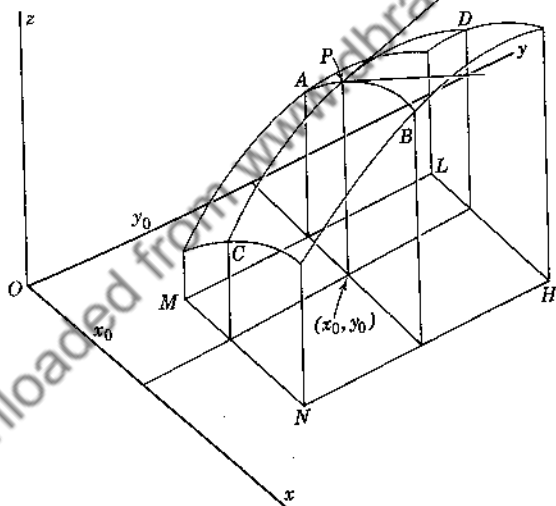


FIG. 3

2. Partial Derivatives. Figure 3 shows the surface representing the functional relation $z = f(x, y)$ over the rectangular region $MNHL$. If (x_0, y_0) is a point in this region, we can cut the surface by the planes $y = y_0$ and $x = x_0$, the intersections being arcs APB and CPD , respectively. We wish to find the slopes of the tangents to the curves of intersection, APB and CPD , at P .

On the curve APB , $z = f(x, y)$ and $y = y_0$; hence, $z = f(x, y_0)$. In this relation, z is an ordinary function of one variable x , and the slope of the tangent at P is the derivative, $\lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$. This

derivative can be found by formula, directly from the expression $f(x, y)$, by first fixing y at y_0 , differentiating the resulting function $f(x, y_0)$, of x alone, by formulas already available, and, after the differentiation, putting $x = x_0$. Since $f(x, y)$ is a function of two variables x and y , and only x is allowed to vary in this differentiation, the result is called the *partial derivative of $f(x, y)$ with respect to x* . It is denoted by $\frac{\partial}{\partial x} f(x, y)$, or $f_x(x, y)$; or more precisely, calling attention to the point (x_0, y_0) , by $\left(\frac{\partial z}{\partial x}\right)_{(x_0, y_0)}$, $\frac{\partial}{\partial x} f(x_0, y_0)$, or $f_x(x_0, y_0)$.

In a similar way, the curve CPD has the equations $z = f(x, y)$, $x = x_0$, whence $z = f(x_0, y)$ along the curve. The slope of this curve at P is given by the derivative of this function of y , $f(x_0, y)$ at $y = y_0$, i.e., $\lim_{k \rightarrow 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}$. This derivative is called the *partial derivative of $f(x, y)$ with respect to y* , and can be formed by formula, first fixing

x at x_0 , differentiating the resulting function of y , and then putting $y = y_0$. This partial derivative is denoted by $\frac{\partial z}{\partial y}$, $\frac{\partial}{\partial y} f(x, y)$, or $f_y(x, y)$; or, more precisely, by $\left(\frac{\partial z}{\partial y}\right)_{(x_0, y_0)}$, $\frac{\partial}{\partial y} f(x_0, y_0)$, or $f_y(x_0, y_0)$.

In order to distinguish the derivative of a function of one variable from a partial derivative, the former is called an *ordinary derivative*.

When obtaining one of these partial derivatives, say f_x , by formula, it is immaterial whether or not we actually replace y by a constant y_0 at the start, as long as we treat y as a constant while we differentiate with respect to x . Some examples will make this clear.

EXAMPLE 1. Find $\partial z / \partial x$ and $\partial z / \partial y$ at $x = 2$, $y = 3$, if

$$z = \frac{1}{4}x^2 + \frac{1}{9}y^2.$$

SOLUTION. We think of y as constant and differentiate

$$z = \frac{1}{4}x^2 + \frac{1}{9}(\text{constant})^2$$

with respect to x , obtaining $\partial z / \partial x = \frac{1}{2}x$. At $x = 2$, $\partial z / \partial x = \frac{1}{2} \cdot 2 = 1$. Similarly, imagining x to be constant, we differentiate

$$z = \frac{1}{4}(\text{constant})^2 + \frac{1}{9}y^2$$

with respect to y , obtaining $\partial z / \partial y = \frac{2}{9}y$. At $y = 3$, $\partial z / \partial y = \frac{2}{9} \cdot 3 = \frac{2}{3}$.

If $y_0 = 3$ and $x_0 = 2$, respectively, are inserted, the work is not essentially changed. This is done in Exercise 2.

EXAMPLE 2. If the surface $z = \frac{1}{4}x^2 + \frac{1}{9}y^2$ is cut by the planes $x = 2$ and $y = 3$, find the slopes of the tangent lines of the resulting curves at the point $(2, 3, 2)$. Hence, find direction cosines of these lines and their equations.

SOLUTION. Along the curve of intersection of $z = \frac{1}{4}x^2 + \frac{1}{9}y^2$ and $y = 3$, z is given by the equation $z = \frac{1}{4}x^2 + 1$, and the slope of this curve is given by $m = \partial z / \partial x = \frac{1}{2}x$. At $x = 2$, $m = \frac{1}{2} \cdot 2 = 1$. This line is perpendicular to the y -axis since it lies in the plane $y = 3$. Hence, $\cos \beta = 0$, and α and γ are complementary. Since $\tan \alpha = 1$, we have $\cos \alpha = \cos \gamma = \sqrt{2}/2$. The equations of the tangent line can now be written in symmetric form: $\frac{x-2}{1} = \frac{y-3}{0} = \frac{z-2}{1}$.

Along the curve of intersection of $z = \frac{1}{4}x^2 + \frac{1}{9}y^2$ and $x = 2$, z is given by the equation $z = 1 + \frac{1}{9}y^2$, and the slope of this curve is given by $m = \partial z / \partial y = \frac{2}{9}y$. At $y = 3$, $m = \frac{2}{9} \cdot 3 = \frac{2}{3}$. This line is perpendicular to the x -axis since it lies in the plane $x = 2$. Hence $\cos \alpha = 0$, and β and γ are complementary. Since, in this case, $\tan \beta = \frac{2}{3}$, we have $\cos \beta = 3/\sqrt{13}$ and $\sin \beta = \cos \gamma = 2/\sqrt{13}$. The equations of this tangent line can now be written: $\frac{x-2}{0} = \frac{y-3}{3} = \frac{z-2}{2}$.

EXAMPLE 3. Find expressions for the partial derivatives with respect to x and y of the function $z = \log(x^2 + y^2) - \sin xy$.

SOLUTION. Thinking of y as constant, we differentiate with respect to x , obtaining

$$\frac{\partial z}{\partial x} = \frac{2x}{x^2 + y^2} - y \cos xy.$$

Similarly,

$$\frac{\partial z}{\partial y} = \frac{2y}{x^2 + y^2} - x \cos xy.$$

EXAMPLE 4. Find $f_x(3, 0)$ and $f_y(3, 4)$ if

$$f(x, y) = (x^2 + y^2)^{\frac{1}{2}} - 2xy + y/x.$$

SOLUTION. We first write

$$\frac{\partial f}{\partial x} = f_x(x, y) = \frac{x}{(x^2 + y^2)^{\frac{1}{2}}} - 2y - \frac{y}{x^2} \quad (2.1)$$

$$\frac{\partial f}{\partial y} = f_y(x, y) = \frac{y}{(x^2 + y^2)^{\frac{1}{2}}} - 2x + \frac{1}{x}. \quad (2.2)$$

Putting $x = 3$, $y = 0$, in (2.1),

$$f_x(3, 0) = \frac{3}{(9 + 0)^{\frac{1}{2}}} - 2 \cdot 0 = \frac{0}{9} = 1.$$

Putting $x = 3$, $y = 4$ in (2.2),

$$f_y(3, 4) = \frac{4}{(9 + 16)^{\frac{1}{2}}} - 2 \cdot 3 + \frac{1}{3} = \frac{73}{15}.$$

The extension of partial differentiation to functions of more variables offers no difficulties, as Example 5 shows.

EXAMPLE 5. Find expressions for $\partial u/\partial x$, $\partial u/\partial y$, and $\partial u/\partial z$ if $u = 3xy^2z^3 + e^x \log(y^2 + z^2)$.

SOLUTION. In order to find $\partial u/\partial x$ we think of y and z as constant. Then

$$\frac{\partial u}{\partial x} = 3y^2z^3 + e^x \log(y^2 + z^2).$$

Similarly, imagining x and z to be constant,

$$\frac{\partial u}{\partial y} = 6xyz^3 + e^x \frac{2y}{y^2 + z^2}.$$

Imagining x and y to be constant,

$$\frac{\partial u}{\partial z} = 9xy^2z^2 + e^x \frac{2z}{y^2 + z^2}.$$

Partial derivatives of higher orders are obtained in the same way as those of the first order. There are three second-order partial derivatives of a function of two variables, and the number increases as we increase the order.

EXAMPLE 6. Find $\frac{\partial^2 z}{\partial x^2}$, $\frac{\partial^2 z}{\partial y^2}$, and $\frac{\partial^2 z}{\partial x \partial y}$, at $(2, 1)$ if

$$z = x^3y - 6x^2y^3.$$

SOLUTION. First, we obtain $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$,

$$\frac{\partial z}{\partial x} = 3x^2y - 12xy^3$$

$$\frac{\partial z}{\partial y} = x^3 - 18x^2y^2.$$

Now,

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \right) = 6xy - 12y^3$$

$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial y} \right) = -36x^2y$$

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y} \right) = 3x^2 - 36xy^2.$$

It is easily verified that $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$, since

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} \right) = 3x^2 - 36xy^2.$$

The "mixed derivatives" $\frac{\partial^2 z}{\partial x \partial y}$ and $\frac{\partial^2 z}{\partial y \partial x}$ are equal, in general, provided both exist and are continuous, although we shall not prove this theorem.

Now, at (2, 1),

$$\frac{\partial^2 z}{\partial x^2} = 6 \cdot 2 \cdot 1 - 12 \cdot 1^3 = 0,$$

$$\frac{\partial^2 z}{\partial y^2} = -36 \cdot 2^2 \cdot 1 = -144,$$

$$\frac{\partial^2 z}{\partial x \partial y} = 3 \cdot 2^2 - 36 \cdot 2 \cdot 1 = -60.$$

The following notation for higher partial derivatives is often used:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx}, \quad \frac{\partial^2 f}{\partial y^2} = f_{yy}, \quad \frac{\partial^2 f}{\partial x \partial y} = f_{xy}, \text{ etc.}$$

EXERCISES

Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ at the points specified in Exercises 1-10.

1. $z = \sqrt{x^2 + y^2}$; (0, 1), (4, -3), (-5, 12), (2, 0).
2. $z = \sin(3x + 5y)$; (0, $\pi/15$), ($\pi/18$, 0), ($\pi/2$, π).
3. $z = x^3 + 3x^2y - y^3 + 10$; (1, 1), (-1, 2), (0, 0), (-3, -1).
4. $z = x/y$; (1, 1), (0, 2), (4, 2), (-3, 6).
5. $z = \sqrt{x^2 + y^2}/x$; (1, 0), (-1, 0), (-3, -4), (4, 3).

6. $z = \sin 2x \cos 3y$; $(0, 0)$, $(\pi/2, \pi/2)$, $(\pi/4, \pi/3)$, $(\pi/6, \pi/4)$.
7. $z = \log(xy/\sqrt{x^2 + y^2})$; $(1, 1)$, $(\frac{1}{2}, \frac{1}{2})$, $(\frac{1}{3}, \frac{1}{3})$, $(2, 4)$.
8. $z = \log \cos xy$; $(1, \pi/4)$, $(\pi/4, -1)$, $(\sqrt{\pi}/2, \sqrt{\pi}/2)$.
9. $z = \tan^{-1}(y/x)$; $(2, 0)$, $(1, 1)$, $(-3, 4)$.
10. $z = (x^2 - y^2)/(x^2 + y^2)$; $(-1, 0)$, $(\frac{1}{2}, \frac{1}{2})$, $(5, 2)$.
11. The sphere $x^2 + y^2 + z^2 = 50$ is cut by the plane $x = 3$. Find the slope of the circle of intersection at $(3, 4, 5)$.
12. If the sphere in Exercise 11 is cut by the plane $y = 4$, find the slope of the circle of intersection at $(-3, 4, 5)$.
13. Reasoning as in Example 2, show that the direction cosines of the tangent to the curve $z = f(x, y)$, $y = y_0$, are proportional to $1, 0, f_x$; similarly, the direction cosines of the tangent to the curve $z = f(x, y)$, $x = x_0$, are proportional to $0, 1, f_y$. Hence, show that equations of these tangents at the point (x_0, y_0, z_0) are, respectively,
- $$\frac{x - x_0}{1} = \frac{y - y_0}{0} = \frac{z - z_0}{f_x}$$
- and
- $$\frac{x - x_0}{0} = \frac{y - y_0}{1} = \frac{z - z_0}{f_y}.$$
- [Hint: For $z = f(x, y)$, $y = y_0$, the slope of the tangent is f_x , and this is equal to $\tan \alpha$. Since $\beta = \pi/2$, $\cos^2 \alpha + \cos^2 \gamma = 1$. Thus, $\cos \alpha = \frac{1}{\sqrt{1 + f_x^2}}$, $\cos \gamma = \frac{f_x}{\sqrt{1 + f_x^2}}$.]
14. Use the results of Exercise 13 to find equations of the tangent to the circle in Exercise 11 at $(3, 4, 5)$. Do the same for the tangent to the circle in Exercise 12 at $(-3, 4, 5)$.

Find $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, and $\frac{\partial u}{\partial z}$ at the points specified in Exercises 15–18.

15. $u = 1/x - y/z$; $(1, 2, -1)$, $(-2, 0, -3)$.
16. $u = \log(x^2 + y^2 + z^2)$; $(1, 1, 1)$, $(0, -1, 2)$.
17. $u = xyz e^{x^2 + y^2 - z^2}$; $(2, -1, 1)$.
18. $u = y \sin xz$; $(1, 2, \pi/3)$, $(\pi/2, 3, -2)$.

Find $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial y^2}$, and $\frac{\partial^2 u}{\partial x \partial y}$ in Exercises 19–22.

19. $u = x/y$.
20. $u = x^3 y^4$.
21. $u = x \sin y$.
22. $u = xye^{xy}$.

23. Show that if $u = \log(x^2 + y^2)$, $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$.

24. If $z = \frac{ax + by}{cx + dy}$, show that $x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 0$.

25. If $u = x \tan^{-1} \frac{y}{x}$, show that $x \frac{\partial u}{\partial x} + y \frac{\partial u}{\partial y} + z \frac{\partial u}{\partial z} = u$.

Exercises 23 and 24 are examples of Euler's theorem on homogeneous functions. This theorem, stated for a function of three variables, says that if $f(x, y, z)$ is homogeneous of order n , then $x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} + z \frac{\partial f}{\partial z} = nf$. The function $f(x, y, z)$ is defined to be homogeneous of order n if the result of replacing the variables x, y, z by tx, ty, tz , respectively, is t^n multiplied by the original function, i.e.,

$$f(tx, ty, tz) = t^n f(x, y, z).$$

In Exercise 23, $n = 0$; in Exercise 24, $n = 1$. See, also, Exercises 16, 17, Art. 8.

3. Tangent Plane to a Surface. The characteristic property of a tangent line to a curve, that we intuitively attach to a tangent plane to a surface, is that of approximating the curve in the neighborhood of the point of tangency more closely than any other line. The student should review Art. 4, Chapter 12, on this point. We recall that a curve has a tangent at (x_0, y_0) if $\frac{\Delta y - dy}{\Delta x} \rightarrow 0$ as $\Delta x \rightarrow 0$. If the equation of the curve is $y = f(x)$, this condition becomes

$$\frac{y - y_0}{x - x_0} = f'(x_0) + \varepsilon,$$

$$y - y_0 = f'(x_0)(x - x_0) + \varepsilon(x - x_0),$$

or

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + \varepsilon(x - x_0), \quad (3.1)$$

in which $\varepsilon \rightarrow 0$ as $(x - x_0) \rightarrow 0$.

We think of a tangent plane as touching a surface much as a tangent line touches a curve. We require it to approximate the surface $z = f(x, y)$ in a neighborhood of a point of tangency (x_0, y_0, z_0) , more closely than any other plane. We shall find, in Art. 4, that such an approximating plane exists, if, and only if, the function $f(x, y)$ satisfies a condition analogous to (3.1). (This condition is stated explicitly in Eq. 4.7.) Let us suppose that $f(x, y)$ meets this condition, and has a tangent plane at (x_0, y_0, z_0) . If both the surface and the tangent plane were cut by any other plane through this point, the line of intersection with the tangent plane must be a tangent line to the curve of intersection with the surface. In particular, if the surface $z = f(x, y)$ is cut by the plane $y = y_0$, the tangent plane must contain the tangent to the curve of intersection. It must also contain the tangent to the curve $z = f(x, y)$, $x = x_0$. Hence, if a tangent

plane exists at (x_0, y_0, z_0) , it must be the plane determined by these two tangent lines, the discussion of which led to partial derivatives, in Art. 2.

The slope of the tangent to the curve $z = f(x, y)$, $y = y_0$, is f_x . Since this line is perpendicular to the y -axis, $\cos \beta = 0$, and we must have $\alpha + \gamma = \pi/2$. Moreover, $\tan \alpha = f_x$. Therefore, $\cos \alpha = 1/\sqrt{1+f_x^2}$ and $\sin \alpha = \cos \gamma = f_x/\sqrt{1+f_x^2}$. That is, $\cos \alpha : \cos \beta : \cos \gamma = 1 : 0 : f_x$. Similarly, the direction cosines of the tangent to the curve $z = f(x, y)$, $x = x_0$, are proportional to $0, 1, f_y$ (see Exercise 13, Art. 2). The direction cosines of the normal to these tangent lines, and hence to the tangent plane, are then, as in Example 5, Art. 12, Chapter 23, proportional to

$$\begin{vmatrix} 0 & f_x \\ 1 & f_y \end{vmatrix}, \quad \begin{vmatrix} f_x & 1 \\ f_y & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \text{ i.e., to } f_x f_y, -1.$$

Hence, the equation of the tangent plane is

$$(x - x_0)f_x + (y - y_0)f_y - (z - z_0) = 0. \quad (3.2)$$

Here, of course, f_x and f_y are to be evaluated at (x_0, y_0) , i.e., these quantities are $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$.

Equation 3.2 is not to be taken as a definition of the tangent plane, since the plane that it represents may fail to meet the approximation requirement laid down for a tangent plane. If a tangent plane exists, however, (3.2) must be its equation. The existence of the partial derivatives f_x and f_y , that appear in Eq. 3.2, is necessary but not sufficient to insure a tangent plane. The conical surface with the equation $z = \sqrt{x^2 + y^2}$, for example, has no true tangent plane at $(0, 0, 0)$, even though f_x and f_y exist at this point and are the slopes of the lines of intersection of the surface with $y = 0$ and $x = 0$, respectively.

4. Existence of a Tangent Plane. The equation of a plane through (x_0, y_0, z_0) is

$$z = z_0 + A(x - x_0) + B(y - y_0). \quad (4.1)$$

If this plane were used to approximate the surface $z = f(x, y)$, containing the point (x_0, y_0, z_0) , the error incurred would be

$$\text{error} = f(x, y) - f(x_0, y_0) - A(x - x_0) - B(y - y_0).$$

Clearly, this error $\rightarrow 0$ as $(x, y) \rightarrow (x_0, y_0)$. But if (4.1) is the equation of a plane that approximates the surface in the same way that a tangent line approximates a curve, as indicated by (3.1), i.e., if (4.1) is the equation of the tangent plane, the error must tend to 0, with $x - x_0$ and $y - y_0$.

more rapidly than for any other plane. This will occur if A and B are so chosen that

$$f(x, y) - f(x_0, y_0) - A(x - x_0) - B(y - y_0) = \varepsilon(x - x_0) + \varepsilon'(y - y_0), \quad (4.2)$$

where ε and ε' both tend to 0 as $x - x_0$ and $y - y_0$ tend to 0 in any way. The condition that two numbers A and B exist so that (4.2) holds is the condition for the existence of the tangent plane.

Assuming this condition to be satisfied, we can determine A and B as follows. Replacing $x - x_0$ by h and $y - y_0$ by k , and transposing, (4.2) becomes

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = Ah + Bk + \varepsilon h + \varepsilon' k. \quad (4.3)$$

Putting $k = 0$, dividing by h , and then letting $h \rightarrow 0$, we find

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = f_x(x_0, y_0) = A. \quad (4.4)$$

Similarly, putting $h = 0$, dividing by k , and then letting $k \rightarrow 0$, we obtain

$$\lim_{k \rightarrow 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k} = f_y(x_0, y_0) = B. \quad (4.5)$$

Substituting these values of A and B in (4.1), we find the equation of the tangent plane to be

$$z = z_0 + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0), \quad (4.6)$$

or

$$f_x \cdot (x - x_0) + f_y \cdot (y - y_0) - (z - z_0) = 0,$$

which agrees with Eq. 3.2.

From (4.6) we see that the direction cosines of the normal to the tangent plane at (x_0, y_0, z_0) , and hence, to the surface at this point, are proportional to $f_x, f_y, -1$.

Inserting in (4.3) the values of A and B from (4.4) and (4.5), we get

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = f_x(x_0, y_0) \cdot h + f_y(x_0, y_0) \cdot k + \varepsilon h + \varepsilon' k, \quad (4.7)$$

in which $\varepsilon \rightarrow 0$, $\varepsilon' \rightarrow 0$ as $h \rightarrow 0$ and $k \rightarrow 0$. The possibility of expressing a function, in a neighborhood of a point (x_0, y_0) , by Eq. 4.7 expresses the condition, equivalent to (4.2), for the surface $z = f(x, y)$ to possess a tangent plane at the point. A function satisfying this condition is said to be *differentiable*. From the above discussion it is clear that a

differentiable function $f(x, y)$ possesses both partial derivatives, f_x and f_y . The converse is not true. For example, for the cone mentioned in Art. 3,

$$z = f(x, y) = \sqrt{x^2 + y^2},$$

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{\sqrt{h^2}}{h} = 1, \quad f_y(0, 0) = \lim_{k \rightarrow 0} \frac{\sqrt{k^2}}{k} = 1.$$

Now,

$$f(0 + h, 0 + k) - f(0, 0) = \sqrt{h^2 + k^2}$$

and

$$f_x(0, 0) \cdot h + f_y(0, 0) \cdot k + \epsilon h + \epsilon' k = h + k + \epsilon h + \epsilon' k.$$

These values do not satisfy (4.7) with ϵ and ϵ' tending to 0 with h and k ; $\sqrt{h^2 + k^2} = h + k + \epsilon h + \epsilon' k$. For, if we let h and k tend to 0 together,

with $h = k$, the ratio $\frac{\sqrt{h^2 + k^2}}{h} = \sqrt{2}$; and

$$\frac{h + k + \epsilon h + \epsilon' k}{h} = 2 + \epsilon + \epsilon'.$$

The situation was otherwise with a function of one variable, which is differentiable if the derivative exists. In the case of two variables, more is needed. Although we shall not prove it, the continuity of the partial derivatives is sufficient to insure differentiability.

EXAMPLE 1. Find the equation of the tangent plane to the paraboloid $z = 30 - x^2 - 2y^2$ at the point $(2, 3, 8)$. Also, find equations of the normal line.

SOLUTION. Differentiating,

$$\partial z / \partial x = -2x, \quad \partial z / \partial y = -4y.$$

At $(2, 3, 8)$,

$$\partial z / \partial x = -4 \quad \text{and} \quad \partial z / \partial y = -12.$$

Hence, using (4.6), the equation of the tangent plane is

$$z = 8 - 4(x - 2) - 12(y - 3),$$

or

$$4x + 12y + z = 52.$$

Since the direction cosines of the normal are proportional to 4, 12, 1, this line has the equations

$$\frac{x - 2}{4} = \frac{y - 3}{12} = z - 8.$$

EXERCISES

Find the equation of the tangent plane, and equations of the normal, at the points specified, in Exercises 1-8.

- $5z = x^2 + y^2$; (3, 4, 5).
- $z = xy$; (1, 1, 1), (2, -1, -2), (0, 0, 0).
- $z = 15 - 4x^2 - y^2$; (1, 3, 2).
- $z = -\sqrt{125 - x^2 - y^2}$; (4, 3, 10), (-4, -3, 10).
- $z = \sqrt{125 - x^2 - y^2}$; (4, 3, -10), (-4, -3, -10).
- $z = 2y/(x + y)$; (1, 1, 1), (2, -4, 4).
- $9x^2 + 16y^2 - 100z = 0$; (-2, -2, 1).
- $z = e^{-(x^2 + 4y^2)}$; (1, $\frac{1}{2}$, e^{-3}).
- Show that the tangent plane to the sphere $x^2 + y^2 + z^2 = a^2$, at (x_0, y_0, z_0) , is $x_0x + y_0y + z_0z = a^2$.
- Show that the tangent plane to the ellipsoid $ax^2 + by^2 + cz^2 = 1$, at (x_0, y_0, z_0) , is $ax_0x + by_0y + cz_0z = 1$.
- Show that at any point common to the two paraboloids, $2z = -1 + x^2 + y^2$ and $2z = 1 - x^2 - y^2$, the normals are perpendicular. (If the normals to two surfaces are perpendicular at a point of intersection, the surfaces are said to intersect *orthogonally* at that point.)
- Show that the two spheres, $x^2 + y^2 + z^2 = 1$ and $(x - 1)^2 + (y - 1)^2 + z^2 = 1$, intersect orthogonally. [Hint: Solve each equation for z . Direction numbers of the normals are $-x/z$, $-y/z$, -1 and $-(x - 1)/z$, $-(y - 1)/z$, -1 , where (x, y, z) satisfy equations of both spheres.]

5. Derivative in a Given Direction. The partial derivative f_x gives the rate of rise of the surface $z = f(x, y)$, proceeding in a direction perpendicular to the y -axis, along a curve formed by cutting the surface by a plane $y = y_0$. It is a rate with respect to x only, and is naturally looked upon as a derivative in the x -direction. Similarly, the partial derivative f_y is a derivative in the y -direction. Both of the tangent lines at a point (x_0, y_0, z_0) , corresponding to these derivatives, must lie in the tangent plane, if there is one. So must tangent lines to curves formed by cutting the surface through this point by any other plane perpendicular to the xy -plane.

Let us suppose that the condition (4.7) for the existence of a tangent plane to $z = f(x, y)$ at (x_0, y_0, z_0) is satisfied. We consider a curve in which this surface is cut by a plane through (x_0, y_0, z_0) , perpendicular to the xy -plane along the line $x = x_0 + r \cos \alpha$, $y = y_0 + r \sin \alpha$, where r is a parameter and α is constant. Then, from (4.7),

$$\begin{aligned} f(x_0 + r \cos \alpha, y_0 + r \sin \alpha) &= f(x_0, y_0) \\ &= f_x \cdot r \cos \alpha + f_y \cdot r \sin \alpha + \epsilon r \cos \alpha + \epsilon' r \sin \alpha. \end{aligned}$$

If we divide by r and let $h = r \cos \alpha$ and $k = r \sin \alpha$, both tend to 0, i.e., $r \rightarrow 0$, we have

$$\lim_{r \rightarrow 0} \frac{f(x_0 + r \cos \alpha, y_0 + r \sin \alpha) - f(x_0, y_0)}{r} = f_x \cos \alpha + f_y \sin \alpha,$$

since $\varepsilon \rightarrow 0$ and $\varepsilon' \rightarrow 0$ as $r \rightarrow 0$. We call this limit the derivative of $f(x, y)$ at (x_0, y_0) in the direction α , and denote it by $D_{(\alpha)}f(x_0, y_0)$. Thus,

$$D_{(\alpha)}f(x_0, y_0) = f_x(x_0, y_0) \cos \alpha + f_y(x_0, y_0) \sin \alpha. \quad (5.1)$$

This directional derivative is clearly the slope of the tangent to the curve of intersection of the surface $z = f(x, y)$ and the plane perpendicular to the xy -plane along the line through (x_0, y_0) in the direction α . The tangent to this curve at (x_0, y_0, z_0) automatically lies in the tangent plane at that point. This is further verified, however, by differentiating the equation of the tangent plane in the direction α :

$$z - z_0 = (x - x_0)f_x + (y - y_0)f_y.$$

In the direction α , $x - x_0 = r \cos \alpha$ and $y - y_0 = r \sin \alpha$. That is,

$$z - z_0 = r \cos \alpha \cdot f_x + r \sin \alpha \cdot f_y.$$

Differentiating with respect to the single variable r ,

$$dz/dr = f_x \cos \alpha + f_y \sin \alpha,$$

a result which agrees with (5.1).

The close relationships connecting the problem of approximating a surface by a plane, the concept of differentiability, the tangent plane, and directional derivatives, are brought out by the following summarizing statements.

1. A function $f(x, y)$ is said to be differentiable if it satisfies (4.7).
2. A surface $z = f(x, y)$ possesses a tangent plane at (x_0, y_0, z_0) if, and only if, $f(x, y)$ is differentiable there.
3. The tangent plane at (x_0, y_0, z_0) approximates the surface in the neighborhood of this point more closely than any other plane.
4. A function $f(x, y)$ possesses a directional derivative at (x_0, y_0) in every direction if, and only if, it is differentiable there.
5. The tangent plane at (x_0, y_0, z_0) is the locus of tangent lines to all curves of intersection of the surface by planes through this point and perpendicular to the xy -plane.

EXAMPLE. Find $D_{(\alpha)}f(x, y)$, if $f(x, y) = \frac{1}{4}x^2 + \frac{1}{9}y^2$, at $(2, 3, 2)$, for $\alpha = 45^\circ$; 30° ; $\tan^{-1} \frac{2}{3}$; $\tan^{-1}(-\frac{3}{2})$; and interpret geometrically.

SOLUTION. Using (5.1), $D_{(\alpha)}f(x, y) = \frac{1}{4}x \cos \alpha + \frac{2}{9}y \sin \alpha$. At $x = 2$, $y = 3$,

$$D_{(\alpha)}f(2, 3) = \cos \alpha + \frac{2}{3} \sin \alpha.$$

For $\alpha = 45^\circ$, $\cos \alpha = \sin \alpha = \sqrt{2}/2$,

$$D_{(\alpha)}f(2, 3) = \frac{\sqrt{2}}{2} + \frac{2}{3} \cdot \frac{\sqrt{2}}{2} = \frac{5}{6} \sqrt{2}.$$

For $\alpha = 30^\circ$, $\cos \alpha = \sqrt{3}/2$, $\sin \alpha = \frac{1}{2}$,

$$D_{(\alpha)}f(2, 3) = \frac{\sqrt{3}}{2} + \frac{1}{3}.$$

For $\alpha = \tan^{-1} \frac{2}{3}$, $\cos \alpha = 3/\sqrt{13}$, $\sin \alpha = 2/\sqrt{13}$,

$$D_{(\alpha)}f(2, 3) = \frac{3}{\sqrt{13}} + \frac{4}{3\sqrt{13}} = \frac{\sqrt{13}}{3}.$$

For $\alpha = \tan^{-1}(-\frac{3}{2})$, $\cos \alpha = 2/\sqrt{13}$, $\sin \alpha = -3/\sqrt{13}$,

$$D_{(\alpha)}f(2, 3) = \frac{2}{\sqrt{13}} - \frac{2}{\sqrt{13}} = 0.$$

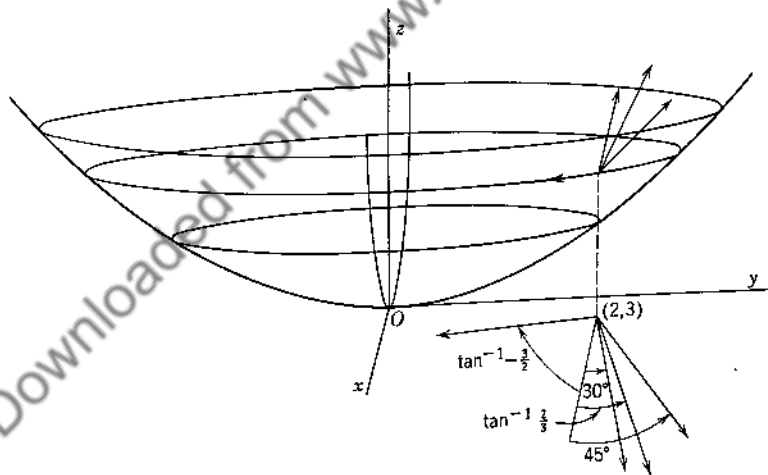


FIG. 4

Each of these derivatives represents the rate of rise on the surface $z = \frac{1}{4}x^2 + \frac{1}{9}y^2$ at $x = 2$, $y = 3$, $z = 2$, when proceeding along a path on the surface lying in the vertical plane inclined at the angle α to the x -axis (see Fig. 4). The last result indicates that in the direction, $\tan^{-1}(-\frac{3}{2})$,

z is not changing, i.e., $\frac{1}{4}x^2 + \frac{1}{9}y^2 = \text{constant}$. Differentiating this implicit relation, we can easily verify the fact that the direction of the tangent to this curve, at $(2, 3)$, in the horizontal plane $z = 2$, is $\tan^{-1}(-\frac{3}{2})$. The direction $\alpha = \tan^{-1}\frac{2}{3}$ is perpendicular to $\tan^{-1}(-\frac{3}{2})$, and, hence, is normal to the curve $\frac{1}{4}x^2 + \frac{1}{9}y^2 = 2$ at $(2, 3)$. The derivative $D_{(\alpha)}$ takes on its maximum value when α has this direction. Thus, differentiating the function of α ,

$$D_{(\alpha)}f(2, 3) = \cos \alpha + \frac{2}{3} \sin \alpha,$$

we get

$$\frac{d}{d\alpha}(\cos \alpha + \frac{2}{3} \sin \alpha) = -\sin \alpha + \frac{2}{3} \cos \alpha.$$

Setting this result equal to 0, in order to determine the value of α corresponding to a maximum or minimum of $D_{(\alpha)}f$, we find $\alpha = \tan^{-1}\frac{2}{3}$. If α has the principal value of $\tan^{-1}\frac{2}{3}$, $D_{(\alpha)}$ has its maximum. If the value of α differs from this by π radians, $D_{(\alpha)}$ has its minimum. These results illustrate a general fact: $D_{(\alpha)}f(x, y)$ is 0 if α is the direction of the tangent to a curve $f(x, y) = \text{constant}$, and has its maximum in the direction of the normal to this curve. This normal derivative is called the *gradient* of $f(x, y)$.

EXERCISES

In each of Exercises 1–3, find the directional derivative at the given point, in the direction indicated.

1. $f(x, y) = x^2 + y^2$; $(2, 3)$; $\alpha = 60^\circ$.
2. $f(x, y) = x^2 - y^2$; $(4, 1)$; $\alpha = 45^\circ$.
3. $f(x, y) = e^{xy}$; $(-1, 2)$; $\alpha = 30^\circ$.
4. Show that $D_{(\alpha)}(x^2 + y^2) = 2\sqrt{x_1^2 + y_1^2}$ at any point (x_1, y_1) , if $\alpha = \tan^{-1}(y_1/x_1)$. Hence, $D_{(\alpha)}$ is constant on the circle $x^2 + y^2 = \text{const.}$ Show, also, that this direction is normal to the circle, and, hence, that the gradient of $x^2 + y^2$ is constant on the circle.
5. As in Exercise 4, show that the gradient of $1/(x^2 + y^2)$ is constant on the circle $x^2 + y^2 = \text{constant}$.
6. If the electric potential V , at any point of a plane, is given by $V = \log \sqrt{x^2 + y^2}$, find the rate of change of potential at any point: (a) in the direction toward the origin; (b) in a direction perpendicular to that toward the origin.

6. Maxima and Minima of Functions of Two Variables. A maximum point of a surface $z = f(x, y)$, like a maximum point of a curve, is one such that all points in a neighborhood of this point are below it, i.e., (x_0, y_0, z_0) is a maximum point if $f(x_0, y_0) - f(x, y)$ is positive for all values of (x, y) in a neighborhood of (x_0, y_0) . It follows that every curve through the point made by the intersection of $z = f(x, y)$ with a plane

perpendicular to the xy -plane must have a maximum there; in particular, the curves $z = f(x, y)$, $y = y_0$, and $z = f(x, y)$, $x = x_0$. Hence, if $f(x, y)$ is differentiable, so that the surface has a tangent plane, this plane must be $z = z_0$, we must have

$$f_x(x_0, y_0) = f_y(x_0, y_0) = 0, \quad (6.1)$$

and all points of the surface near (x_0, y_0, z_0) must lie below this plane. Similar statements apply to a minimum point, all points of the surface near this point lying above the tangent plane.

As with curves, maxima and minima may occur at the boundary of the region over which the surface is described, at which points condition 6.1 need not apply. Of course, this condition may not apply at a point where $f(x, y)$ is not differentiable, e.g., the apex of a cone.

Although (6.1) is a necessary condition for a maximum or minimum of a differentiable function, this condition is not sufficient to guarantee that (x_0, y_0, z_0) is such a point. This is easily seen from the equation $z = xy$, which represents an hyperbolic paraboloid (Art. 8, Chapter 23). This surface is saddle shaped, rising from the origin for values of x and y that agree in sign, i.e., (x, y) in the first or third quadrant, and falling for values of x and y of different signs, (x, y) in the second or fourth quadrant. At the origin, however, condition (6.1) is satisfied, and the tangent plane is the xy -plane itself.

In applications, it is often possible to locate maxima and minima by means of the equations

$$f_x(x, y) = 0, \quad f_y(x, y) = 0,$$

and to determine whether such a point is actually a maximum or a minimum or neither, by comparison with values of $f(x, y)$ for nearby values of (x, y) . More extended treatment of the problem will be found in advanced textbooks.

The above statements carry over with little change to functions of more than two variables. Thus, for example, in order that $x = x_0$, $y = y_0$, $z = z_0$ correspond to a maximum or a minimum of the differentiable function $f(x, y, z)$, in the interior of a region of definition of the function, it is necessary that

$$f_x(x_0, y_0, z_0) = f_y(x_0, y_0, z_0) = f_z(x_0, y_0, z_0) = 0. \quad (6.2)$$

EXAMPLE 1. Examine $f(x, y) = x^2 + y^2 + 4$ for maxima and minima.
SOLUTION. Differentiating,

$$f_x = 2x, \quad f_y = 2y.$$

Setting $f_x = f_y = 0$, we find $x = y = 0$. Clearly, $(0, 0)$ makes $x^2 + y^2 + 4$ a minimum.

EXAMPLE 2. Find the shape of a rectangular tank with open top to contain a given volume, in order that the total area shall be least.

SOLUTION. Denoting the length, width, and depth by x , y , and z , respectively, the function to be minimized is

$$S = xy + 2xz + 2yz. \quad (6.3)$$

Since the volume, $V = xyz$, is fixed, we can substitute $z = V/xy$ in (6.3), obtaining S as a function of x and y only.

$$S = xy + 2V(x + y)/xy = xy + 2V/y + 2V/x.$$

Differentiating,

$$\partial S/\partial x = y - 2V/x^2, \quad \partial S/\partial y = x - 2V/y^2.$$

Setting $\partial S/\partial x = \partial S/\partial y = 0$, we have

$$y - 2V/x^2 = 0, \quad x - 2V/y^2 = 0,$$

or

$$x^2y = 2V, \quad y^2x = 2V, \quad \text{whence} \quad x = y = \sqrt[3]{2V}.$$

Substituting this result in $z = V/xy$, we get $z = \frac{1}{2}\sqrt[3]{2V}$. Hence, the tank has a square base and the depth is $\frac{1}{2}$ the length.

One can be convinced intuitively that these results correspond to a minimum as follows. For one dimension very small, the total area is large and decreases gradually as this dimension is increased, until a minimum is reached, after which it increases again. That is, a minimum exists, and since $\partial S/\partial x = \partial S/\partial y = 0$ at a minimum, the results obtained by use of this condition correspond to a minimum.

EXERCISES

1. Find the minimum value of the function $x^2/16 + y^2/4 - 25$.
2. Find the maximum value of the function $z = \sqrt{100 - (x - 2)^2 - (y + 5)^2}$.
3. Show that the surface $z = x^2 - y^2$ has no maximum or minimum point.
4. Find a point at which the function $f(x, y) = x^2 + xy + y^2 - 3x - 6y$ satisfies conditions (6.1). Show that it is a minimum point by comparing $f(x_0, y_0)$ with $f(x_0 + \varepsilon, y_0 + \varepsilon')$, where ε and ε' are any two numbers such that $\varepsilon^2 + \varepsilon'^2 > 0$.
5. Find the shortest distance from the origin to the plane $x + \sqrt{2}y - z = 12$. Check by the result of Exercise 9, Art. 11, Chapter 23.
6. If $f(x, y) = \sin x \sin y$, show that $f_x = f_y = 0$ at $(\pi/2, \pi/2)$ and at (π, π) . Show, further, that $(\pi/2, \pi/2)$ corresponds to a maximum, but that (π, π) corresponds to neither a maximum nor a minimum.
7. If $f(x, y) = \sin x + \sin y$, show that $x = y = \pi/2$ corresponds to a maximum, and $x = y = 3\pi/2$ to a minimum; but that $x = \pi/2, y = 3\pi/2$ produces neither, although $f_x = f_y = 0$ at this point.

8. Divide the positive number a into three parts such that their product is maximum.
9. Show that the rectangular parallelepiped of given volume and minimum surface is a cube.
10. Find the points on the surface $xyz = 1$ that are nearest the origin.
11. Two lines are given by the parametric equations

$$x = -2 - 3t, y = 5 + 2t, z = 3 + t$$

$$x = 1 + s, y = -8 + 6s, z = 10 - 2s.$$

Find the shortest distance between them. Show that the shortest segment joining them is perpendicular to both lines. [Hint:

$$(\text{distance})^2 = \{(-2 + 3t) - (1 + s)\}^2 + \text{etc.}]$$

12. A building is made in the form of a cylinder surmounted by a conical roof. If the radius = r , the height of the cylinder = h , and the height of the cone = a , show that the shape that will enclose a given volume with a minimum surface is $a : h : r = 2 : 1 : \sqrt{5}$, the area of the base not to be considered.

7. Differential of a Function of Several Variables. The attempt to approximate a function of one variable in the neighborhood of a point by a linear function led to the notion of a differential. It was shown that the tangent was the best possible linear approximation to the curve $y = f(x)$ near a point (x_0, y_0) . The change in y on the curve, resulting from the change Δx in x was denoted by Δy , and that on the tangent corresponding to a change dx in x , was denoted by dy . In this way, two new variables were introduced, connected by the relation $dy = f'(x_0) dx$. If dx and Δx are made to agree, then dy is the best linear approximation to Δy , and, hence, $y_0 + dy$ is the best linear approximation to values of y on the curve in a neighborhood of (x_0, y_0) .

In the same way, it can be shown that the best possible linear approximation to a differentiable function of two variables, $z = f(x, y)$, is afforded by the tangent plane. Denoting by Δz the change in z resulting from changes of Δx and Δy in the independent variables, we introduce the variable dz to represent the change in z on the tangent plane, corresponding to changes of dx and dy in the independent variables. Thus, three new variables are introduced, connected by the relation

$$dz = f_x(x_0, y_0) dx + f_y(x_0, y_0) dy. \quad (7.1)$$

If dx and dy are identified with Δx and Δy , respectively, then dz is the best linear approximation to Δz . This can be shown by the method of Art. 4, Chapter 12. It also follows immediately from the definition of differentiability (4.7).

Equation 7.1 is identical with the equation of the tangent plane

$$z - z_0 = f_x \cdot (x - x_0) + f_y \cdot (y - y_0). \quad (4.6)$$

On the other hand, the equation

$$\Delta z = f_x \Delta x + f_y \Delta y + \varepsilon \Delta x + \varepsilon' \Delta y, \quad (7.2)$$

or

$$z - z_0 = f_x \cdot (x - x_0) + f_y \cdot (y - y_0) + \varepsilon(x - x_0) + \varepsilon'(y - y_0),$$

is equivalent to Eq. 4.3 of the surface itself (see Fig. 5). The actual *difference*, Δz , of the function, that results from the variation from (x_0, y_0) to (x, y) , is given by (7.2). The linear approximation dz to this

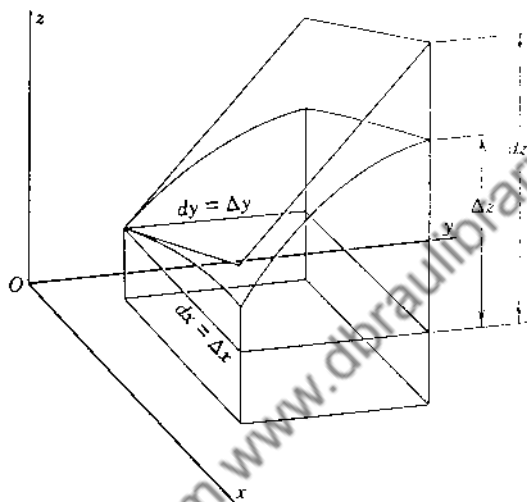


FIG. 5

difference, given by Eq. 7.1, is called the *differential* of the function. This quantity is also called the *total differential*, and $f_x dx + f_y dy$ is called the *exact differential* of the function $f(x, y)$.

A differential of a function of more than two independent variables is defined in the same way. That is, if $u = f(x, y, z)$,

$$du = f_x dx + f_y dy + f_z dz,$$

in which dx , dy , and dz are new independent variables often identified with Δx , Δy , and Δz , respectively.

We shall illustrate the use of differentials in approximations.

EXAMPLE 1. A board of green lumber is 12 ft long, 10 in. wide, and 1 in. thick. As it dries, its length remains constant, but its width decreases by .15 in. and its thickness by .02 in. What is the approximate decrease in volume?

SOLUTION. We have $V = 12 \cdot 12 wt$, in cu in. Forming the differential dV as an approximation to ΔV ,

$$dV = 144 w dt + 144 t dw.$$

Setting $t = 1$, $w = 10$, $dt = -.02$, $dw = -.15$, we have

$$dw = 144 \cdot 10(-.02) + 144 \cdot 1(-.15) = -50.4.$$

The shrinkage is therefore approximately 50.4 cu in.

EXAMPLE 2. Two sides and included angle of a triangle are found by measurement to be $a = 12$ in., $b = 16$ in., and $C = 60^\circ$. If the errors in measurement are not more than .01 in. on each side and $30'$ in the angle, estimate the maximum error in the area.

SOLUTION. The formula for the area is $A = \frac{1}{2}ab \sin C$. Forming the differential of A as an approximation to ΔA ,

$$dA = \frac{1}{2}ab \cdot \cos C \cdot dC + \frac{1}{2}b \cdot \sin C \cdot da + \frac{1}{2}a \cdot \sin C \cdot db.$$

Putting $a = 12$, $b = 16$, $C = 60^\circ$, $dC = .00873$ (rad), $da = .01$, $db = .01$, we have

$$\begin{aligned} dA &= \frac{1}{2} \left[12 \cdot 16 \cdot \frac{1}{2} (.00873) + 16 \cdot \frac{\sqrt{3}}{2} (.01) + 12 \cdot \frac{\sqrt{3}}{2} (.01) \right] \\ &= .54, \text{ correct to two decimal places.} \end{aligned}$$

EXERCISES

Find the expression for du in each of Exercises 1-12.

1. $u = 3x^2y + 2xy^2$.

2. $u = xyz$.

3. $u = \log(xy)$.

4. $u = x^4 - x^2y^2 + y^4 + 2$.

5. $u = xe^{yz}$.

6. $u = xe^{y+z}$.

7. $u = e^{-x} \sin y$.

8. $u = y/x$.

9. $u = \tan^{-1}(y/x)$.

10. $u = (x^2 + y^2)/z$.

11. $u = \sqrt{x^2 + y^2}$.

12. $u = \log \sqrt{x^2 + y^2}$.

13. A circular cylinder has a radius of 10 in. and is 4 ft long. Approximate the increase in volume if the radius increases by .5 in. and the length by 1 in.

14. Estimate the error in computing the third side of the triangle in Example 2 by the law of cosines.

15. The period, T seconds, of a simple pendulum is given by $T = 2\pi\sqrt{l/g}$, in which l is the length in feet and g is the acceleration of gravity in feet per second per second. In an experiment, l and g are found to be 3.26 and 32.2, respectively, with maximum errors of .02 in each figure. Find T corresponding to these measurements, and estimate the maximum error in the calculated result. [Take dl and dg opposite in sign to produce the maximum error.]

16. Find the approximate relative error in Exercise 15. [Hint: Take the logarithms of both sides.]

17. Find an approximate value of $\sqrt{(16.02)(48.95)}$ by the use of differentials. [Hint: Find $u + du$.]

18. Find an approximate value of $(2.07) \log 2.8$ by the use of differentials.

8. Compound Functions. The Chain Rule. Partial differentiation is carried out by the same formulas that are used for ordinary differentiation. With compound functions, however, in which two or more intermediate variables are involved, the chain rule of Chapter 3 is obviously inappropriate. If there is one intermediate variable and more final variables, the rule can be used without change. From Chapter 3, if

$$u = g(x) \quad \text{and} \quad x = \phi(t),$$

then

$$du/dt = g'(x) \cdot \phi'(t).$$

It follows immediately, from the definition of a partial derivative, that if

$$u = g(x) \quad \text{and} \quad x = \phi(t, s),$$

then

$$\left. \begin{aligned} \partial u / \partial t &= g'(x) \cdot \phi_t(t) \\ \partial u / \partial s &= g'(x) \cdot \phi_s(s) \end{aligned} \right\} \quad (8.1)$$

Now, let us consider a function of a single final variable t through two intermediate variables x and y . That is,

$$u = g(x, y), \quad x = \phi(t), \quad y = \psi(t),$$

so that $u = f(t)$. We assume that $g(x, y)$, $\phi(t)$, and $\psi(t)$ are all differentiable functions. Corresponding to an increment Δt , there are then the quantities Δx , Δy , and Δu which tend to 0 with Δt , and which satisfy Eq. 4.7.

$$\Delta u = g_x \Delta x + g_y \Delta y + \varepsilon \Delta x + \varepsilon' \Delta y,$$

where ε and ε' tend to 0 with Δx and Δy . Dividing by Δt ,

$$\frac{\Delta u}{\Delta t} = g_x \frac{\Delta x}{\Delta t} + g_y \frac{\Delta y}{\Delta t} + \varepsilon \frac{\Delta x}{\Delta t} + \varepsilon' \frac{\Delta y}{\Delta t}.$$

Letting $\Delta t \rightarrow 0$, we have

$$du/dt = f'(t) = g_x(x, y)\phi'(t) + g_y(x, y)\psi'(t). \quad (8.2)$$

or, in the Leibniz notation,

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt}. \quad (8.3)$$

Equations 8.2 and 8.3 express the chain rule for a function of two variables, each of which is a function of a single variable. If t is

identical with x , i.e., $u = g(x, y)$, where $y = \psi(x)$, formulas 8.2 and 8.3 become

$$f'(x) = g_x(x, y) + g_y(x, y)\psi'(x), \quad (8.4)$$

and

$$\frac{du}{dx} = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} \frac{dy}{dx} \quad (8.5)$$

If there are two final variables instead of one, say $u = g(x, y)$, $x = \phi(t, s)$, and $y = \psi(t, s)$, so that $u = f(t, s)$, formula 8.2 is replaced by two formulas for $f_t(t, s)$ and $f_s(t, s)$, as follows.

$$\left. \begin{aligned} f_t(t, s) &= g_x(x, y)\phi_t(t, s) + g_y(x, y)\psi_t(t, s) \\ f_s(t, s) &= g_x(x, y)\phi_s(t, s) + g_y(x, y)\psi_s(t, s) \end{aligned} \right\} \quad (8.6)$$

The student can easily write (8.6) in Leibniz notation. Again, if one of the final variables is equal to one of the intermediate variables, say $t = x$, one of Eqs. 8.6 takes a special form. Suppose $z = g(x, y)$, and $y = \psi(x, s)$ so that $z = f(x, s)$. Then (8.6) becomes

$$\left. \begin{aligned} f_x &= g_x \cdot 1 + g_y \psi_x = g_x + g_y \psi_x \\ f_s &= g_x \cdot 0 + g_y \psi_s = g_y \psi_s \end{aligned} \right\} \quad (8.7)$$

Equations 8.7 are superior in form to the same equations written in Leibniz notation, viz.,

$$\left. \begin{aligned} \frac{\partial z}{\partial x} &= \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial x} \\ \frac{\partial z}{\partial s} &= \frac{\partial z}{\partial y} \frac{\partial y}{\partial s} \end{aligned} \right\} \quad (8.8)$$

In the first of Eqs. 8.8 the symbol $\frac{\partial z}{\partial x}$ does not mean the same on both sides of the equation. On the right, it denotes the derivative of $g(x, y)$ with respect to x , on the assumption that y is fixed, i.e., x and s vary so as to keep $y = \psi(x, s)$ constant. On the left, the same symbol denotes the derivative of $z = f(x, y)$ with respect to x , on the assumption that s (and not y) is fixed. The notations in (8.7) avoid this confusion.

The foregoing discussion suggests that it is important to know what variables are held constant during a partial differentiation, rather than to give sole attention to the variable that changes. For example, suppose

$$z = x + y, \quad x = t, \quad y = s.$$

Then,

$$\partial z / \partial t = 1, \quad \partial z / \partial s = 1.$$

But if $z = x + y$, $x = t$, $y = r - 2t$,

$$\partial z / \partial t = 1, \quad \partial z / \partial r = 1.$$

Thus, $\partial z / \partial t$ is equal to 1 in one case and -1 in the other, although t has the same meaning in both. We see, however, that the variables held constant while obtaining $\partial z / \partial t$ differ in the two instances. In the first we hold $s = y$ constant, and in the second we hold $r = y + 2x$ constant.

Further extensions of the chain rule to deal with more variables, either more intermediate variables or more final variables, offer no additional difficulties. Thus, for example, if

$$u = g(x, y, z), \quad \text{where } x = \phi(t, s), \quad y = \psi(t, s), \quad z = \chi(t, s),$$

$$\partial u / \partial t = g_x \phi_t + g_y \psi_t + g_z \chi_t$$

$$\partial u / \partial s = g_x \phi_s + g_y \psi_s + g_z \chi_s.$$

By means of the chain rule, we can prove an important theorem on differentials. It will be recalled from Chapter 12 that the formula for the differential of a function of one variable, $y = f(x)$, viz., $dy = f'(x) dx$, holds whether x is the independent variable or is itself a function of a third variable. A similar statement applies to differentials of functions of several variables. Thus, suppose $u = f(t, s)$ through the compound function

$$u = g(x, y), \quad x = \phi(t, s), \quad y = \psi(t, s).$$

Now, in the first place, $du = f_t dt + f_s ds$. Operating through the compound relation, we have

$$\text{and} \quad f_t = g_x \phi_t + g_y \psi_t$$

$$\text{Hence,} \quad f_s = g_x \phi_s + g_y \psi_s.$$

$$\begin{aligned} du &= (g_x \phi_t + g_y \psi_t) dt + (g_x \phi_s + g_y \psi_s) ds \\ &= g_x (\phi_t dt + \phi_s ds) + g_y (\psi_t dt + \psi_s ds) \\ &= g_x dx + g_y dy. \end{aligned}$$

The proof for more variables follows in the same way.

EXAMPLE 1. If $z = \sin xy$ and $x = e^{-t}$, $y = \sin^{-1} t$, find dz/dt .

SOLUTION. By formula 8.3, we have

$$dz/dt = y \cos xy (-e^{-t}) + (x \cos xy) / \sqrt{1-t^2}.$$

The result can be put in terms of t alone by substituting $x = e^{-t}$, $y = \sin^{-1} t$.

EXAMPLE 2. If $u = e^{\sin^2 x + \cos x}$, find du/dx .

SOLUTION. We write $u = e^{y+z}$, $y = \sin^2 x$, $z = \cos x$. Applying (8.3),

$$\begin{aligned} du/dx &= e^{y+z} dy/dx + e^{y+z} dz/dx \\ &= e^{\sin^2 x + \cos x} (2 \sin x \cos x - \sin x). \end{aligned}$$

EXAMPLE 3. If $u = x^2 + xy + y^2$, $x = r \cos \theta$, $y = r \sin \theta$, find $\partial u/\partial r$ and $\partial u/\partial \theta$.

SOLUTION. Using (8.6),

$$\begin{aligned} \partial u/\partial r &= (2x + y) \cos \theta + (x + 2y) \sin \theta \\ \partial u/\partial \theta &= (2x + y)r(-\sin \theta) + (x + 2y)r \cos \theta. \end{aligned}$$

EXAMPLE 4. If $z = \sqrt{x} + \sin xy$ and $y = \tan^{-1} x$, find dz/dx .

SOLUTION. Using (8.5),

$$dz/dx = 1/(2\sqrt{x}) + y \cos xy + (x \cos xy)/(1 + x^2).$$

EXERCISES

1. If $z = \log(x^2 - y^2)$, $x = \sin t$, $y = e^t$, find dz/dt .
2. If $u = x \log x$, $x = \sin\left(t^2 - \frac{1}{s}\right)$, find $\frac{\partial u}{\partial t}$ and $\frac{\partial u}{\partial s}$.
3. If $u = \tan^{-1} x$, $x = p^2 q - \sin q$, find $\partial u/\partial p$ and $\partial u/\partial q$.
4. If $z = x^2 - xy - y^2$, $x = 2s - t$, $y = s - 2t$, find $\partial z/\partial t$ and $\partial z/\partial s$.
5. If $u = \sin^{-1}\left(\frac{t-s}{t+s}\right)$, find $\partial u/\partial t$ and $\partial u/\partial s$. [Let $t - s = x$, $t + s = y$.]
6. If $z = r^2 e^{y/x}$, $x = r \cos \theta$, $y = r \sin \theta$, find $\partial z/\partial r$ and $\partial z/\partial \theta$.
7. If $u = \sin \sqrt{x^2 + y^2}$, $x = t + 2s$, $y = t^2 - s^2$, find $\partial u/\partial t$ and $\partial u/\partial s$.
8. If $u = (x - y)/(1 + xy)$, $x = \tan t$, $y = \tan s$, find $\partial u/\partial t$ and $\partial u/\partial s$ in two ways.
9. If $z = r^2 e^\theta$, $r = \sqrt{x^2 + y^2}$, $\theta = \tan^{-1}(y/x)$, find $\partial z/\partial x$ and $\partial z/\partial y$.
10. If $u = \tan^{-1}(y + ax) - \log(y - ax)$, find $\partial u/\partial x$ and $\partial u/\partial y$.
11. If $z = f(xy)$, show that $x \frac{\partial u}{\partial x} = y \frac{\partial u}{\partial y}$.
12. If $z = f(x, y)$, and $x = t + s$, $y = t - s$, show that $\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = \frac{\partial^2 z}{\partial t \partial s}$.
13. Show that $a^2 \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2}$, if $u = f(x + at) + g(x - at)$.
14. If $u = f(x, y)$, and $x = r \cos \theta$, $y = r \sin \theta$, find $\frac{\partial u}{\partial r}$ and $\frac{\partial u}{\partial \theta}$, and show that $\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2} \left(\frac{\partial u}{\partial \theta}\right)^2$.

15. Prove that if $z = f\left(\frac{y}{x}\right)$, $x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = 0$.

16. Prove that if $z = x^n f\left(\frac{y}{x}\right)$, $x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = nz$. [This is a form of Euler's theorem on homogeneous functions, see Exercise 17.]

17. Prove Euler's theorem on homogeneous functions, for a function of two variables, viz., if $f(x, y)$ is homogeneous of order n , so that $f(tx, ty) = t^n f(x, y)$, then $xf_x + yf_y = nf$. (See Exercises 24, 25, Art. 2.) [Hint: Differentiate $f(tx, ty) = t^n f(x, y)$ with respect to t , and then let $t = 1$.]

18. If $u = g(x, y)$ and $x = \phi(t)$, $y = \psi(t)$, show that

$$\frac{d^2 u}{dt^2} = g_{xx}(\phi')^2 + 2g_{xy}\phi'\psi' + g_{yy}(\psi')^2 + g_x\phi'' + g_y\psi''.$$

[Hint: Find $\frac{d}{dt}\left(\frac{du}{dt}\right) = \frac{d}{dt}[g_x(x, y)\phi'(t) + g_y(x, y)\psi'(t)]$, being careful to note that $\frac{d}{dt}g_x(x, y) = g_{xx}\phi' + g_{xy}\psi'$, etc.]

19. If $u = g(x, y)$ and $x = \phi(t, s)$, $y = \psi(t, s)$, show that

$$\frac{\partial^2 u}{\partial t^2} = g_{xx}(\phi_t)^2 + 2g_{xy}\phi_t\psi_t + g_{yy}(\psi_t)^2 + g_x\phi_{tt} + g_y\psi_{tt}.$$

20. If u and v are functions of x and y , satisfying the equations $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$, $\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$, show that $\frac{\partial u}{\partial r} = \frac{1}{r} \frac{\partial v}{\partial \theta}$, $\frac{1}{r} \frac{\partial u}{\partial \theta} = -\frac{\partial v}{\partial r}$, where $x = r \cos \theta$, $y = r \sin \theta$.

9. Implicit Functions. The differentiation of a function of one variable, defined by an implicit relation, was taken up intuitively in Chapter 3, Art. 13. Here we shall generalize and deepen that discussion to some extent, although without supplying complete proofs.*

We shall consider y as a function of x , defined by the implicit relation

$$f(x, y) = 0. \quad (9.1)$$

Now, in the first instance, we cannot be sure that (9.1) actually defines y as a function of x . Some conditions must be imposed upon $f(x, y)$ in order for this to be true, and especially for the function so defined to be continuous and differentiable. Sufficient conditions for this are given by the following theorem.

THEOREM 1. If $f(x, y)$ is defined in some region including a point (x_0, y_0) at which $f(x_0, y_0) = 0$, and possesses partial derivatives which are continuous throughout the region, with $f_y(x_0, y_0) \neq 0$, then there is some neighborhood of (x_0, y_0) within which the equation $f(x, y) = 0$ determines

* See Courant, Vol. II, Chapter III.

y as a continuous, differentiable function of x : $y = \phi(x)$, with $y_0 = \phi(x_0)$, and $y' = -f_x/f_y$.

Of course, it may not be possible to write $\phi(x)$ explicitly. Although we shall not prove this theorem in its entirety, it is easy to establish the formula for the derivative, if the existence of the continuous differentiable function $\phi(x)$ is assumed. When we think of Eq. 9.1 as defining y as a function of x , we regard it as an identity in x . That is, upon insertion of $y = \phi(x)$, this equation becomes $f[x, \phi(x)] \equiv 0$. With this in mind, differentiating (9.1) with respect to x , we obtain, by use of (8.4),

$$f_x + f_y y' = 0.$$

Solving for y' ,

$$y' = -f_x/f_y. \quad (9.2)$$

EXAMPLE 1. Show that $x^2 + y^2 - 25 = 0$ defines y as a continuous, differentiable function of x in a neighborhood including the point $(3, 4)$, by reference to Theorem 1, and find y' at this point.

SOLUTION. The function $f(x, y) = x^2 + y^2 - 25$ is a continuous function of x and y in a neighborhood of $(3, 4)$, where

$$f(3, 4) = 3^2 + 4^2 - 25 = 0.$$

Moreover, $f_x = 2x$ and $f_y = 2y$, which clearly exist, and $f_y \neq 0$ in a neighborhood including no point of the x -axis, such as the rectangular area in Fig. 6. Applying (9.2), we have

$$y' = -f_x/f_y = -2x/2y = -\frac{3}{4}.$$

An equation in three variables of the form

$$F(x, y, z) = 0 \quad (9.3)$$

defines one variable implicitly, as a function of the other two, if the function $F(x, y, z)$ satisfies conditions analogous to those in Theorem 1.

THEOREM 2. If $F(x, y, z)$ is defined in a region including (x_0, y_0, z_0) , and $F(x_0, y_0, z_0) = 0$; and if $F(x, y, z)$ possesses continuous partial derivatives F_x , F_y , and F_z throughout the region, with $F_z(x_0, y_0, z_0) \neq 0$; then, within some neighborhood of (x_0, y_0, z_0) , Eq. 9.3 defines z as a continuous, differentiable function of x and y : $z = \phi(x, y)$, with $z_0 = \phi(x_0, y_0)$, and

$$\phi_x = -F_x/F_z, \quad \phi_y = -F_y/F_z. \quad (9.4)$$

EXAMPLE 2. Show that the equation $x + y + z - \sin(xyz) = 0$ defines z as a function of x and y in a neighborhood of $(0, 0, 0)$, by reference to Theorem 2, and find $\partial z/\partial x$ at this point.

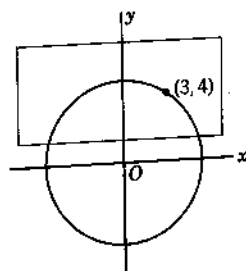


FIG. 6

SOLUTION. The function, $F = x + y + z - \sin(xyz)$, is equal to 0 at $(0, 0, 0)$ and is continuous everywhere. The partial derivatives are $F_x = 1 - yz \cos(xyz)$, $F_y = 1 - xz \cos(xyz)$, $F_z = 1 - xy \cos(xyz)$, and are continuous everywhere; also, $F_z(0, 0, 0) = 1 \neq 0$. Hence, $z = a$ function of x and y : $z = \phi(x, y)$, and, by (9.4),

$$\frac{\partial z}{\partial x} = \phi_x = \frac{1 - yz \cos(xyz)}{1 - xy \cos(xyz)},$$

which is equal to -1 at $(0, 0, 0)$.

EXAMPLE 3. Show that the equation $xy - \log xy - 1 = 0$ does not satisfy the conditions of Theorem 1 so as to define $y = \phi(x)$ in a neighborhood of $(1, 1)$.

SOLUTION. We have $f(x, y) = xy - \log xy - 1$, $f(1, 1) = 1 - 0 - 1 = 0$, $f_x = y - 1/x$, $f_y = x - 1/y$. These partial derivatives are continuous at $(1, 1)$, but $f_y(1, 1) = 0$. Hence, it is not to be expected that the given equation defines y as a function of x in a neighborhood of $(1, 1)$. Theorem 1, as stated, does not make this certain, as it only states sufficient conditions. The condition, $f_y(1, 1) \neq 0$, is necessary, however, and this equation does not define y as a function of x . Neither does it define x as a function of y , for a similar reason.

10. Curves and Surfaces Given in Implicit Form. Implicit equations of curves in two dimensions are very familiar, many of the curves studied in analytic geometry being usually so expressed; e.g., circles, ellipses, and hyperbolas. Equation 9.2 can be used to find the directions of tangents and normals of curves in this form. In general, if a curve is given by the equation

$$f(x, y) = 0, \quad (10.1)$$

in which $f(x, y)$ satisfies the conditions of the theorem in Art. 9, the slope of the tangent is

$$y' = -f_x/f_y.$$

Hence, the equation of the tangent at a point satisfying (10.1) is

$$y - y_0 = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)} \cdot (x - x_0)$$

or

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0. \quad (10.3)$$

The equation of the normal is

$$y - y_0 = \frac{f_y(x_0, y_0)}{f_x(x_0, y_0)} \cdot (x - x_0)$$

or

$$f_y(x_0, y_0)(x - x_0) - f_x(x_0, y_0)(y - y_0) = 0. \quad (10.4)$$

Equations 10.3 and 10.4 are valid in any region in which $f(x, y) = 0$ determines y as a function of x or x as a function of y , i.e., if either f_x or f_y is different from 0.

The second derivative y'' can be found from (10.2).

$$y'' = \frac{d}{dx} \left[-\frac{f_x(x, y)}{f_y(x, y)} \right] = -\frac{f_y(f_{xx} + f_{yx}y') - f_x(f_{xy} + f_{yy}y')}{(f_y)^2}$$

Substituting for y' from (10.2) and simplifying, on the assumption that $f_{yx} = f_{xy}$,

$$y'' = -\frac{f_{xx}f_y^2 - 2f_{xy}f_xf_y + f_{yy}f_x^2}{f_y^3} \quad (10.5)$$

Surfaces were given by implicit equations in Chapter 23, e.g., spheres, ellipsoids, and cones. Let a surface have the equation

$$F(x, y, z) = 0, \quad (10.6)$$

in which $F(x, y, z)$ satisfies the conditions of Theorem 2, Art. 9. Then, $\partial z/\partial x$ and $\partial z/\partial y$ are found by (9.4). Substituting these expressions for $\partial z/\partial x$ and $\partial z/\partial y$ in the equation of the tangent plane, (3.2), we have

$$z - z_0 = -\frac{F_x}{F_z}(x - x_0) - \frac{F_y}{F_z}(y - y_0),$$

or

$$F_x \cdot (x - x_0) + F_y \cdot (y - y_0) + F_z \cdot (z - z_0) = 0. \quad (10.7)$$

As with Eqs. 10.3 and 10.4, the validity of (10.7) does not require $F_z \neq 0$, but it suffices if one of the three partial derivatives, F_x , F_y , or F_z , is different from 0.

The direction cosines of the normal are proportional to F_x , F_y , F_z , and hence the normal line has the equations,

$$\frac{x - x_0}{F_x} = \frac{y - y_0}{F_y} = \frac{z - z_0}{F_z}. \quad (10.8)$$

As a special result, we can state that two surfaces, $F(x, y, z) = 0$ and $G(x, y, z) = 0$, intersect orthogonally at any point where

$$F_x G_x + F_y G_y + F_z G_z = 0. \quad (10.9)$$

EXERCISES

Show that none of the equations in Exercises 1-4 satisfies Theorem 1, Art. 9, in a neighborhood of the origin; but that each defines y as a function of x in a neighborhood of the point specified; and find y' at that point.

1. $x \sin xy = 0$, $(1, \pi)$.
2. $xy + \log x + \log y = 1$, $(1, 1)$.

3. $(x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0$, $(a\sqrt{3}/2, a/2)$.

4. $y \sin(x + y) - xe^{-xy} = 0$, $(\pi/4, \pi/4)$.

Find $\partial z/\partial x$ and $\partial z/\partial y$ from each of the equations in Exercises 3–5.

5. $xyz = a$.

6. $(x + y)(y + z)(z + x) = a$.

7. $x^2 + y^2 + z^2 - \log yz^2 = 0$.

8. $\sin xy + \sin yz + \sin zx = C$.

9. Find the equation of the tangent plane, and equations of the normal line to $xy + yz + zx = 1$, at $(1, 0, 1)$.

10. Find the equation of the tangent plane and equations of the normal line to the sphere $x^2 + y^2 + z^2 = 1$, at any point (x_0, y_0, z_0) on it.

11. Find the equation of the tangent plane, and equations of the normal line to the ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, at any point (x_0, y_0, z_0) on it.

12. Find the equation of the tangent plane to the surface $x^2 + y^2 + z^2 = a^2$, at any point (x_0, y_0, z_0) on it, and show that the sum of the squares of the intercepts on the coordinate axis is the same for all tangent planes.

13. Substitute from Eqs. 10.2 and 10.5 in the formula for curvature,

$$K = \frac{y''}{(1 + y'^2)^{3/2}},$$

to obtain the curvature from an implicit equation.

$$K = \frac{f_{xx}f_y^2 - 2f_{xy}f_xf_y + f_{yy}f_x^2}{(f_x^2 + f_y^2)^{3/2}}.$$

[The denominator can have either sign.]

14. Show that two curves, $f(x, y) = 0$ and $g(x, y) = 0$, intersect orthogonally if, and only if, $f_x g_x + f_y g_y = 0$.

15. Show that two curves, $f(x, y) = 0$ and $g(x, y) = 0$, have a common tangent at the point of intersection if, and only if, $f_x g_y - f_y g_x = 0$ at that point.

16. Show that every parabola of the family $y^2 - 2p(x - p/2) = 0$ for which $p > 0$, intersects every parabola for which $p < 0$, at right angles. [Hint: Consider two values of p of opposite sign, say m and $-n$. For these two curves the product of the slopes is $-mn/y^2$. At a point of intersection, $y^2 = mn$.]

17. Show that the spheres, $x^2 + y^2 + z^2 = 1$ and $(x - a)^2 + (y - b)^2 + (z - c)^2 = r^2$, intersect orthogonally if, and only if, $a^2 + b^2 + c^2 = 1 + r^2$.

18. The surfaces, $x^2 + y^2 + z^2 = 14$ and $xyz = 6$, have the point $(2, 3, 1)$ in common. Find the angle between their normals at that point.

11. Space Curves. Curves in three dimensions were discussed in Chapter 23, Art. 13. There we saw how to find the tangent line and normal plane to a curve expressed parametrically, and the special situation in which the parameter was x or y . We now solve these problems for curves of intersection of surfaces expressed in implicit form.

$$F(x, y, z) = 0, \quad G(x, y, z) = 0. \quad (11.1)$$

In order to find direction numbers of the tangent to this curve at any point (x_0, y_0, z_0) , we note that the normals to both surfaces at this point lie in the normal plane to the curve, which is, in fact, determined by these two lines. The direction cosines of the normals to the surfaces (11.1) are proportional to F_x, F_y, F_z and G_x, G_y, G_z , respectively. Hence, direction numbers a, b, c of the tangent to the curve are

$$a = k \begin{vmatrix} F_y & F_z \\ G_y & G_z \end{vmatrix}, \quad b = k \begin{vmatrix} F_z & F_x \\ G_z & G_x \end{vmatrix}, \quad c = k \begin{vmatrix} F_x & F_y \\ G_x & G_y \end{vmatrix}, \quad (11.2)$$

where k is any number different from 0. This follows from Example 5, Art. 12, Chapter 23. In these expressions, each partial derivative is evaluated at the point (x_0, y_0, z_0) .

We can now write equations of the tangent line at (x_0, y_0, z_0) .

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}. \quad (11.3)$$

The equation of the normal plane at this point is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \quad (11.4)$$

In (11.3) and (11.4), a, b , and c have values obtained from (11.2).

EXAMPLE. Find the tangent line and normal plane to the curve $xyz = 4$, $x^2 + y^2 + z^2 = 9$, at the point $(2, 2, 1)$.

SOLUTION. Putting $F(x, y, z) = xyz - 4$, $G(x, y, z) = x^2 + y^2 + z^2 - 9$, we have

$$F_x = yz, \quad F_y = zx, \quad F_z = xy$$

$$G_x = 2x, \quad G_y = 2y, \quad G_z = 2z.$$

At $(2, 2, 1)$: $F_x = 2$, $F_y = 2$, $F_z = 4$, $G_x = 4$, $G_y = 4$, $G_z = 2$. From (11.2) we have $a = -12k$, $b = 12k$, $c = 0$. Letting $k = -1/12$, $a = 1$, $b = -1$, $c = 0$. The tangent then has the equations

$$\frac{x - 2}{1} = \frac{y - 2}{-1} = \frac{z - 1}{0},$$

or

$$x + y = 4, \quad z = 1.$$

The equation of the normal plane is

$$(x - 2) \cdot 1 + (y - 2)(-1) + (z - 1) \cdot 0 = 0,$$

or

$$x - y = 0.$$

EXERCISES

1. Write equations of the tangent line and normal plane to the curve in the illustrative example, (a) at $(2, -2, -1)$; (b) at any point on it, (x_0, y_0, z_0) .

2. Treating the equations of a curve: $y = f(x)$, $z = g(x)$, as implicit equations, i.e., $F(x, y, z) = y - f(x)$, $G(x, y, z) = z - g(x)$, show that the tangent line has the equations,

$$\frac{x - x_0}{1} = \frac{y - y_0}{f'(x_0)} = \frac{z - z_0}{g'(x_0)}.$$

3. Let a curve be given parametrically by $x = \phi(t)$, $y = \psi(t)$, $z = \chi(t)$. Denoting the inverse function of ϕ by ϕ^{-1} , i.e., $t = \phi^{-1}(x)$, we can write $y = \psi[\phi^{-1}(x)]$, $z = \chi[\phi^{-1}(x)]$. Hence, use the result of Exercise 2 to show that the tangent line has the equations

$$\frac{x - x_0}{\phi'(t_0)} = \frac{y - y_0}{\psi'(t_0)} = \frac{z - z_0}{\chi'(t_0)}.$$

Find equations of the tangent line and the normal plane to each of the curves in Exercises 4–7, at the point specified.

4. $x^2 + y^2 + z^2 = 9$, $x + y + z = 1$, at $(2, 1, -2)$.

5. $x^2 + y^2 + z^2 = 4z + 8$, $2x - 3y + z = 14$, at $(2, -2, 4)$.

6. $36z = 9x^2 + 4y^2$, $2x = z^2$, at $(2, 3, 2)$.

7. $x^2 + y^2 = 50$, $\tan 2z = y/x$, at $(5, 5, \pi/8)$.

8. $x^2 + y^2 + z = 20$, $x^2 + y^2 = 6x$, at $(3, 3, 2)$.

Downloaded from www.dbraulibrary.org

Functions of Several Variables Integration

1. Volume under a Surface. Volumes bounded by curved surfaces were calculated in Chapters 9 and 14. This was possible because the special volumes there considered could be viewed as limits of sums leading to definite integrals. The general problem of calculating the

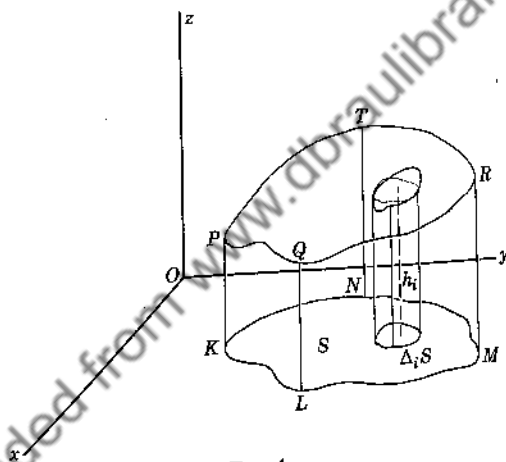


FIG. 1

volume bounded by curved surfaces requires an extension of the concept of a definite integral. The solution of this problem depends upon the solution of the following more restricted problem.

Consider a domain S in the xy -plane, bounded by a closed curve $KLMN$ (Fig. 1). Let $z = f(x, y)$ be a continuous, positive function over this domain. Imagine a cylindrical surface intersecting the xy -plane perpendicularly along $KLMN$ and cutting the surface $z = f(x, y)$ in a space curve $PQRT$. Our problem is to find the volume bounded by these surfaces: the surface $z = f(x, y)$, the cylindrical surface, and the xy -plane.

We approach this problem in the same way that we dealt with the area problem in Chapter 7. First, we divide the domain of the independent

variables in the xy -plane into n subregions (Fig. 2). On each of these subregions as base we erect a cylinder intersecting the surface, and choose for its height some value of the height of the surface over that subregion (see Fig. 1). Clearly, the sum of the volumes of all these cylinders approximates the required volume.

Denoting the area of a typical subregion $\Delta_i S$ by $\Delta_i A$, and the height of the cylinder erected on it by h_i , the volume of the cylinder is $h_i \Delta_i A$.

As an approximation to the required volume, we have

$$V \approx \sum_{i=1}^n h_i \Delta_i A$$

In view of the solution of the area problem, it is natural to try to obtain the limit of this sum as $n \rightarrow \infty$. We must observe a precaution, however, in order to obtain such a limit uniquely, i.e., the same limit for all manners of subdividing the domain and for all choices of h_i in each subdivision. We require that, as $n \rightarrow \infty$, not only must

each $\Delta_i A$ tend to 0, but the greatest distance across each subdivision $\Delta_i S$ must tend to 0. With this precaution, it can be proved that the sum approaches a unique limit, and the volume is defined to be this limit. That is,

$$V = \lim_{n \rightarrow \infty} \sum_{i=1}^n h_i \Delta_i A.$$

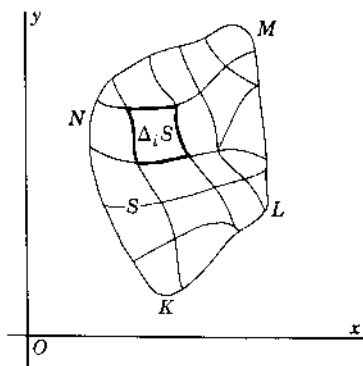
Since h_i = some value of z , say z_i , corresponding to a point (x_i, y_i) in $\Delta_i S$, we write

$$V = \lim_{n \rightarrow \infty} \sum_{i=1}^n z_i \Delta_i A = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i, y_i) \Delta_i A.$$

2. A Double Integral. The sum just expressed, and its limit, are the two-dimensional analogues of the sum $\sum_{i=1}^n f(x_i) \Delta_i x$ and the definite integral $\int_a^b f(x) dx$, discussed in Chapter 7. The factor $\Delta_i A$ is two dimensional, the domain over which summation takes place is two dimensional, and the function $f(x, y)$ contains two independent variables. Under the conditions described in Art. 1, however, viz., $f(x, y)$ continuous over the domain and the greatest diameter of $\Delta_i S$ tending to 0 as n tends to ∞ , it can be proved that a unique limit exists. We shall not give the proof.*

* See Courant, Vol. II, Chapter IV.

FIG. 2



This limit defines the integral of $f(x, y)$ over the domain S , and is called a two-dimensional *domain*, or *double*, integral. Thus, we write

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i, y_i) \Delta_i A = \int_S f(x, y) dA, \quad (2.1)$$

with the understanding that the greatest diameter of $\Delta_i S$ tends to 0 as n tends to ∞ , and the summation extends over all subregions of domain S .

The proof of the existence of this limit shows that it has the same value for all ways of subdividing S , and for all choices of (x_i, y_i) in each $\Delta_i S$. Because of this fact we can, if we wish, make the subdivisions rectangles with dimensions $\Delta_i x$ and $\Delta_i y$, so that $\Delta_i A = \Delta_i x \cdot \Delta_i y$. The integral is, therefore, often written $\int_S f(x, y) dx dy$. It may be that, in this process,

the domain S can only be approximated by rectangular subdivisions, e.g., if the boundary is curved. Even so, when we pass to the limit to obtain the integral, the error due to approximating the domain by rectangular subdivisions tends to 0, so that this mode of subdivision yields the integral.

Several of the theorems on integrals in Chapter 7, Art. 10, can be extended to double integrals without difficulty. We call special attention to the following.

$$\int_S f(x, y) dA = \int_{S_1} f(x, y) dA + \int_{S_2} f(x, y) dA,$$

where S_1 and S_2 are nonoverlapping parts making up the whole domain S . The limitation to positive functions, in Art. 1, was imposed for the geometrical consideration of volume, and is no longer necessary in (2.1), which is numerical in character.

The evaluation of double integrals by direct summation is difficult, except in trivial examples. Even for one-dimensional, or *single*, integrals, we found that evaluation by means of inverse derivatives, which was justified by the fundamental theorem, was far more convenient than direct summation. We need something analogous to an inverse derivative for a double integral. This is supplied by the *iterated integral*, which will be discussed in Art. 3.

3. A Two-Dimensional Iterated Integral. We can extend the integral concept to two dimensions in a way different from that in Art. 2. Let us first suppose the domain S , in the xy -plane, to be a rectangle bounded by the lines $x = a$, $x = b$, $y = c$, and $y = d$, (Fig. 3). Now, we imagine y fixed at some number y_1 between c and d , and form the single integral with respect to x of the function $f(x, y_1)$: $\int_a^b f(x, y_1) dx$. Thus, we perform a partial integration of the function $f(x, y)$, with y held fixed,

analogous to partial differentiation. This integration can be performed with y any value in the interval from c to d , without specifying a particular y , the result being a function of y . We therefore write

$$\int_a^b f(x, y) dx = F(y). \quad (3.1)$$

Equation 3.1 can be interpreted geometrically as follows. The solid bounded by the xy -plane, the four planes $x = a$, $x = b$, $y = c$, and

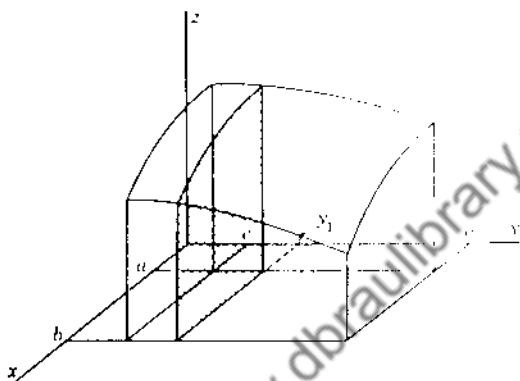


FIG. 3

$y = d$, and the surface $z = f(x, y)$, is intersected by a plane $y = \text{constant}$ (Fig. 3). The integral $\int_a^b f(x, y) dx$ is the area of the intersection. If y is set equal to different values, different areas are to be expected, and, in general, the area obtained is a function of y .

We might now define the volume of the solid, illustrated in Fig. 3, differently from the definition in Art. 1. We imagine the solid to be cut into slices by planes perpendicular to the y -axis.

$$y = y_0 = c, y = y_1, y = y_2, \dots, y = y_n = d.$$

The areas of these plane sections are

$$F(y_0) = \int_a^b f(x, y_0) dx, F(y_1) = \int_a^b f(x, y_1) dx, \dots, F(y_n) = \int_a^b f(x, y_n) dx,$$

respectively. The volume of a typical slice is approximated by the product $F(y_i) \Delta y$, where y_i is some value of y in the i th slice, and Δy is the thickness. Then, $V \approx \sum_{i=1}^n F(y_i) \Delta y$, and it is natural to define the volume of the solid as the limit of the sum as $n \rightarrow \infty$ and each $\Delta y \rightarrow 0$, viz.,

$$V = \int_c^d F(y) dy = \int_c^d \left[\int_a^b f(x, y) dx \right] dy.$$

This integral, also denoted by $\int_c^d dy \int_a^b f(x, y) dx$ and by $\int_c^d \int_a^b f(x, y) dx dy$, is called an *iterated* integral. It is obtained by first integrating $f(x, y)$ with respect to x , as though y were fixed, and then integrating the resulting function of y with respect to y . It thus differs in concept from the double integral, in which we think of the result as being obtained by a single limiting process extending over the whole domain in one step.

Intuitively, it appears that the iterated integral and the double integral must result in the same volume, and this suggests their equality. It is to be emphasized, however, that these two definitions are different, and the equality of the results to which they lead must be proved.* We shall take this equality for granted without proof, and write

$$\int_S f(x, y) dA = \int_c^d dy \int_a^b f(x, y) dx, \quad (3.2)$$

in which S is the rectangle in Fig. 3.

Equation 3.2 furnishes the method we seek for evaluating a double integral over a rectangular domain. Since two integrations are required in this evaluation, it is common to write the double integral with two integral signs, $\iint_S f(x, y) dA$, or $\iint_S f(x, y) dx dy$.

A second iterated integral can be formed, in which we integrate first with respect to y , holding x fixed, and then integrate the resulting function of x . This integral over the above rectangular domain is written

$$\int_a^b dx \int_c^d f(x, y) dy \quad \text{or} \quad \int_a^b \int_c^d f(x, y) dy dx.$$

With rectangular domains, it is immaterial which of these iterated integrals is used to evaluate the double integral, i.e.,

$$\iint_S f(x, y) dA = \int_c^d dy \int_a^b f(x, y) dx = \int_a^b dx \int_c^d f(x, y) dy.$$

The evaluation of double integrals over nonrectangular domains will be explained in Art. 4. We first take up some examples with rectangular domains.

EXAMPLE 1. Find $\iint_S (x^2 + y^2) dA$ over the rectangle bounded by $x = 0$, $x = 5$, $y = 0$, and $y = 2$.

* See Courant, Vol. II, Chapter IV.

SOLUTION. $\iint_S (x^2 + y^2) dA = \int_0^2 dy \int_0^5 (x^2 + y^2) dx$. Integrating $\int_0^5 (x^2 + y^2) dx$ as though y were constant, we obtain

$$\left[\frac{x^3}{3} + xy^2 \right]_0^5 = \frac{125}{3} + 5y^2.$$

Performing the second integration,

$$\int_0^2 \left(\frac{125}{3} + 5y^2 \right) dy = \left[\frac{125}{3} y + \frac{5}{3} y^3 \right]_0^2 = 96\frac{2}{3}.$$

EXAMPLE 2. Find the volume in the first octant, bounded by the xy -plane, the surface $z = xy$, and the planes $x = 1$, $y = 1$ (Fig. 4).

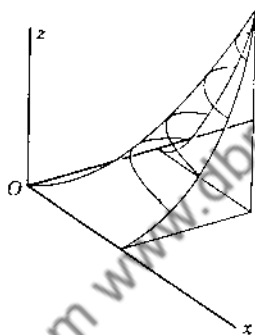


FIG. 4

SOLUTION. This volume is equal to the double integral $\iint_S xy dA$, where S is the rectangle bounded by $x = 0$, $x = 1$, $y = 0$, and $y = 1$. Hence, by (3.2),

$$V = \iint_S xy dA = \int_0^1 dy \int_0^1 xy dx = \int_0^1 \frac{y}{2} dy = \frac{1}{4}.$$

EXERCISES

Evaluate the iterated integrals in Exercises 1–10.

1. $\int_0^4 dy \int_0^4 (x^2 + y^2) dx.$

2. $\int_0^3 dx \int_1^2 (2x^2 + 3xy) dy.$

3. $\int_{-1}^1 dy \int_{-1}^1 (x^2 - 2xy) dx.$

4. $\int_0^2 dy \int_0^2 (10 - 2x - 4y) dx.$

5. $\int_0^{10} dx \int_0^{10} (x^2 - y^2) dy.$

6. $\int_2^4 dx \int_0^1 xe^{xy} dy.$

$$7. \int_0^{\pi/2} dx \int_0^{\pi/2} (\sin x - x \cos y) dy.$$

$$8. \int_0^1 dx \int_0^1 \frac{x}{\sqrt{1-x^2} \sqrt{1-(xy)^2}} dy.$$

$$9. \int_{1/e}^1 dy \int_1^e \frac{\log x}{xy} dx.$$

$$10. \int_4^8 dy \int_2^{10} \left(\frac{1}{x} + \frac{1}{y} \right) dx.$$

Evaluate the double integrals in Exercises 11–15.

$$11. \iint_S (x+y) dA; S \text{ is the square with vertices } (0, 0), (0, 2), (2, 0), (2, 2).$$

$$12. \iint_S (10 - x^2 - y^2) dA; S \text{ is the square in Exercise 11.}$$

$$13. \iint_S (10 - x^2 - y^2) dA; S \text{ is the square with vertices } (1, 1), (1, -1), (-1, 1), (-1, -1).$$

$$14. \iint_S y^2 \sin xy dA; S \text{ is the rectangle with vertices } (0, 0), (\pi/2, 0), (0, 1), (\pi/2, 1).$$

$$15. \iint_S x e^{x+y} dA; S \text{ is the rectangle with vertices } (0, 0), (0, \log 2), (\log 2, 0), (\log 2, \log 2).$$

$$16. \text{ Find the volume of the prismoid bounded by the planes } z = 0, z = x + 2y, x = 1, x = 3, y = 2, y = 5.$$

$$17. \text{ A hole 2 in. square is cut through a circular cylinder 4 in. in diameter, the axis of the cylinder and that of the hole intersecting at right angles. Find the volume of the hole.}$$

$$18. \text{ Find the volume between the } xy\text{-plane and the hyperbolic paraboloid } z = x^2 - y^2, \text{ based on the rectangle with vertices } (1, 1), (1, -1), (3, 1), (3, -1).$$

$$19. \text{ Find the volume in the first octant bounded by the } xy\text{-plane, the surface } z = \frac{xy}{\sqrt{x^2 + y^2}}, \text{ and the planes } x = 1, y = 1.$$

$$20. \text{ Find the volume bounded by the surface } z = 16 - x^2 - y^2 \text{ and the planes } 3x = x + 3y, x = 0, x = 3, y = 0, y = 2.$$

4. Integrals over Nonrectangular Domains. If the domain S of a two-dimensional integral has curved boundaries, the integral is still equal to an iterated integral. Again, consider a function of y defined by an integral with respect to x , viz.,

$$F(y) = \int_{\phi(y)}^{\psi(y)} f(x, y) dx.$$

Here, as in Art. 3, y is to be treated as a constant in the evaluation of the integral. On this account, the limits $\phi(y)$ and $\psi(y)$ are also temporarily constants, and this integral is not essentially different from other definite integrals. The resulting function, $F(y)$, can be integrated over an interval of y -values, say from c to d , obtaining the iterated integral,

$$\int_c^d F(y) dy = \int_c^d \left[\int_{\phi(y)}^{\psi(y)} f(x, y) dx \right] dy = \int_c^d dy \int_{\phi(y)}^{\psi(y)} f(x, y) dx. \quad (4.1)$$

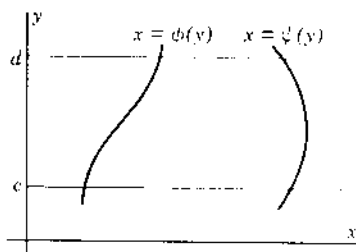


FIG. 5

The geometrical interpretation of this iterated integral is similar to that in Art. 3. Imagine the domain in the xy -plane bounded by the two curves $x = \phi(y)$ and $x = \psi(y)$, and the two lines $y = c$ and $y = d$ (Fig. 5).

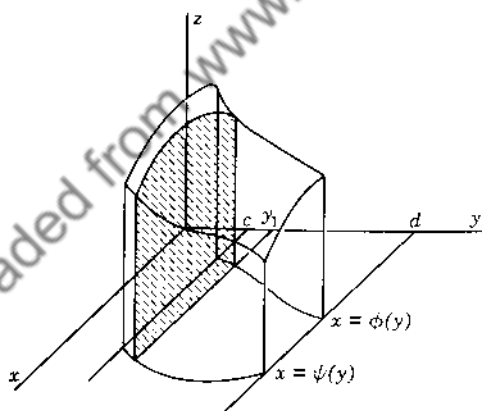


FIG. 6

Over this domain we have the surface $z = f(x, y)$ (Fig. 6). If we fix y at y_1 , and then evaluate $\int_{\phi(y)}^{\psi(y)} f(x, y) dx$, we obtain the area of the plane section of the solid in Fig. 6, at $y = y_1$. By considerations like those brought out in Art. 3, we might be led to define the volume over this domain by the integral $\int_c^d F(y) dy = \int_c^d dy \int_{\phi(y)}^{\psi(y)} f(x, y) dx$, and, hence, to

identify this iterated integral with the double integral $\iint_S f(x, y) dA$ over the same domain. Again, we take this identification for granted, without proof.

If the domain S is bounded by two curves $y = g(x)$, $y = h(x)$, and two lines $x = a$, $x = b$ (Fig. 7), the double integral can be evaluated by the iterated integral

$$\iint_S f(x, y) dA = \int_a^b dx \int_{g(x)}^{h(x)} f(x, y) dy. \quad (4.2)$$

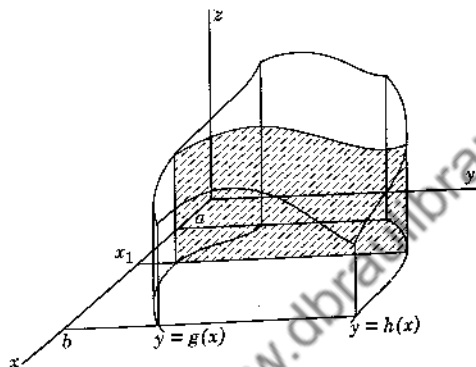


FIG. 7

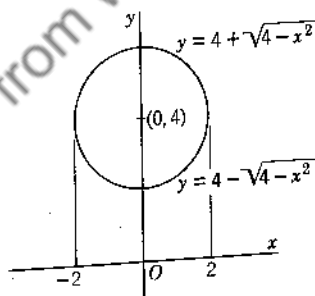


FIG. 8

If S is bounded by a closed curve, either of Eqs. 4.1 or 4.2 can be used, according to convenience. For example, let S be the interior of the circle $x^2 + y^2 - 8y + 12 = 0$ (Fig. 8). Solving this equation for y ,

$$y = 4 \pm \sqrt{4 - x^2}.$$

The upper half of the circle has the equation $y = 4 + \sqrt{4 - x^2}$, which we identify with $y = h(x)$ in Eq. 4.2; the lower half has the equation

$y = 4 - \sqrt{4 - x^2}$, which we identify with $g(x)$. Then

$$\iint_S f(x, y) dA = \int_{-2}^2 dx \int_{4 - \sqrt{4 - x^2}}^{4 + \sqrt{4 - x^2}} f(x, y) dy.$$

A more general situation is illustrated in Fig. 9. The problem of finding expressions for $g(x)$ and $h(x)$, however, may offer practical difficulties.

The determination of the limits of an iterated integral of the form (4.2) is facilitated in practice by conceiving the first integration as corresponding to a summing up along a line parallel to the y -axis, from the lower boundary to the upper, as suggested by MN in Fig. 9.

We must be sure that the functions $g(x)$ and $h(x)$ actually represent the lower and upper boundaries, respectively, for every x between a and b .

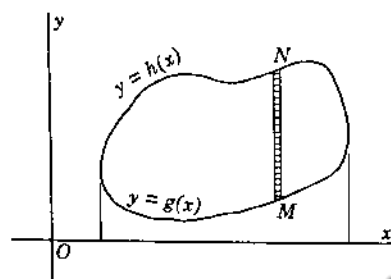


FIG. 9

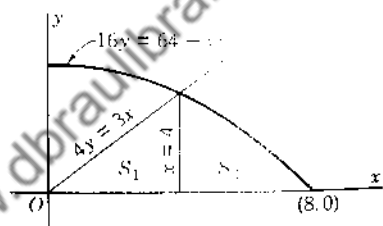


FIG. 10

It often happens that one of these functions cannot be represented by the same expression throughout this range. Thus, for example, if S is the region bounded by the x -axis, the line $4y = 3x$, and the parabola $16y = 64 - x^2$ (Fig. 10), the upper boundary is represented by two expressions $y = \frac{3}{4}x$ for $0 \leq x \leq 4$, and $y = 4 - \frac{1}{16}x^2$ for $4 < x \leq 8$. The use of these two expressions in the appropriate ranges has the effect of breaking S into two parts.

$$\begin{aligned} \iint_S f(x, y) dA &= \iint_{S_1} f(x, y) dA + \iint_{S_2} f(x, y) dA \\ &= \int_0^4 dx \int_0^{\frac{3}{4}x} f(x, y) dy + \int_4^8 dx \int_0^{4 - \frac{1}{16}x^2} f(x, y) dy, \end{aligned}$$

in which S_1 and S_2 are indicated in Fig. 10. A similar discussion applies to iterated integrals of the form (4.1). In the example here given, the breakup into two parts can be avoided by integrating with respect to x first. Example 1 illustrates both methods.

EXAMPLE 1. Evaluate $\iint_S xy dA$ over the domain in Fig. 10.

SOLUTION. Using the iterated integral of (4.2), we write

$$\begin{aligned}\iint_S xy \, dA &= \int_0^4 dx \int_0^{4-x} xy \, dy + \int_4^8 dx \int_4^{4-\frac{1}{2}x^2} xy \, dy \\ &= \int_0^4 \frac{9}{32} x^3 \, dx + \int_4^8 (8x - \frac{1}{4}x^3 + \frac{1}{512}x^5) \, dx \\ &= 18 + 36 = 54.\end{aligned}$$

The iterated integral (4.1) leads to the solution,

$$\begin{aligned}\iint_S xy \, dA &= \int_0^3 dy \int_{\frac{1}{2}y}^{\sqrt{4-y}} xy \, dx \\ &= \int_0^3 (32y - 8y^2 - \frac{8}{9}y^3) \, dy \\ &= 54.\end{aligned}$$

EXAMPLE 2. Evaluate $\iint_S \cos x \, dA$, if S is bounded by the y -axis and the curves $y = \sin x$ and $y = \cos x$ (Fig. 11).

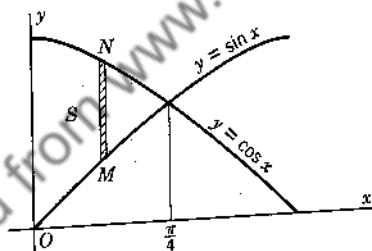


FIG. 11

SOLUTION. We use (4.2). The first integration corresponds to a summing up along any vertical line MN from the lower boundary, $y = \sin x$, to the upper boundary, $y = \cos x$. The second integration then corresponds to a summing up of such "vertical" integrals, from the extreme left to the extreme right, i.e., from $x = 0$ to $x = \pi/4$. Hence,

$$\begin{aligned}\iint_S \cos x \, dA &= \int_0^{\pi/4} dx \int_{\sin x}^{\cos x} \cos x \, dy \\ &= \int_0^{\pi/4} (\cos^2 x - \cos x \sin x) \, dx \\ &= \pi/8.\end{aligned}$$

EXAMPLE 3. Find the volume of the paraboloid $2z = x^2 + y^2$ cut off by the plane $z = y + 1$.

SOLUTION. The elimination of z between the given equations yields an equation connecting x and y on the curve of intersection of the paraboloid by the plane. Its space locus is the cylinder with elements perpendicular to the xy -plane and passing through this curve. In the xy -plane, it is the curve bounding the region S over which the required volume lies (Fig. 12). This equation is $\frac{1}{2}(y^2 + x^2) = y + 1$, or

$$x^2 + (y - 1)^2 = 3. \quad (4.3)$$

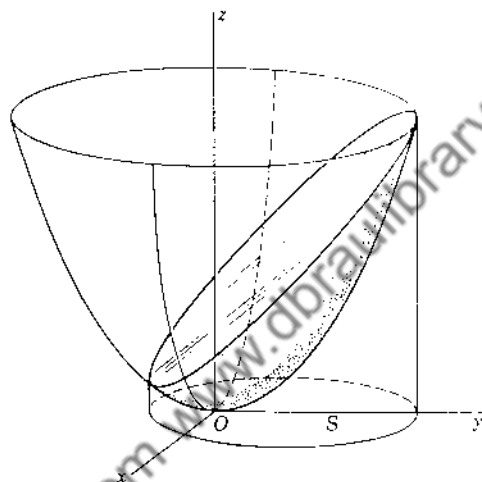


FIG. 12

The required volume is the difference of the cylindrical volumes lying over this circle, one cut off by the plane $z = y + 1$, and the other cut off by the paraboloid $2z = x^2 + y^2$. Hence,

$$\begin{aligned} V &= \iint_S (y + 1) \, dA - \iint_S \frac{1}{2}(x^2 + y^2) \, dA \\ &= \iint_S [(y + 1) - \frac{1}{2}(x^2 + y^2)] \, dA. \end{aligned}$$

Since the solid is symmetrical with respect to the yz -plane, we may find the half on the positive side of this plane and double the result. For this purpose, we set up the iterated integral

$$\frac{V}{2} = \int_{-1-\sqrt{3}}^{1+\sqrt{3}} dy \int_0^{\sqrt{3-(y-1)^2}} (y + 1 - \frac{1}{2}x^2 - \frac{1}{2}y^2) \, dx,$$

obtaining the limits from (4.3). Integrating,

$$\frac{V}{2} = \int_{-\sqrt{3}}^{1-\sqrt{3}} \frac{1}{3} [3 - (y-1)^2]^{\frac{1}{2}} dy = \frac{1}{3} \int_{-\sqrt{3}}^{\sqrt{3}} (3 - u^2)^{\frac{1}{2}} du, \quad u = y - 1,$$

$$= \frac{9}{8}\pi, \quad \text{and} \quad V = \frac{9}{4}\pi.$$

In order to obtain correct limits of an iterated integral by which to evaluate a double integral, it is advisable to make a drawing in the xy -plane, clearly showing the boundaries of S . This will also help to decide which order of integration in the iterated integral is preferable. A space figure is an aid in visualizing a solid whose volume is required. The act of constructing such a figure often assists the imagination more than the figure itself.

EXERCISES

Evaluate the double integrals in Exercises 1-10.

1. $\iint_S xy \, dA$; S bounded by $y = 0$, $x = 4$, $y^2 = x$.
2. $\iint_S \frac{x}{y^2} \, dA$; S bounded by $y^2 = x^2$, $y = 2x$.
3. $\iint_S y\sqrt{x} \, dA$; S bounded by $y^2 + 2x = 4$, $x = 0$.
4. $\iint_S y \, dA$; S bounded by $x^2 + y^2 = 1$.
5. $\iint_S (x + y) \, dA$; S bounded by $y = 0$, $y = x^2$, $y = (x - 1)^2$.
6. $\iint_S (x^2 + xy) \, dA$; S bounded by $x = 0$, $y = x$, $x + y = 2$.
7. $\iint_S \sin x \, dA$; S bounded by $y = x$, $y = 0$, $x = \pi/2$.
8. $\iint_S \frac{x}{x^2 + y^2} \, dA$; S bounded by $y = x$ and $ay = x^2$.
9. $\iint_S e^{x/y} \, dA$; S bounded by $y^2 = x$, $y = 1$, $x = 0$.
10. $\iint_S e^x \, dA$; S bounded by $x + y = 1$, $y = \cosh x$, $x = 1$.

11. Find the volume of the tetrahedron bounded by the coordinate planes and the plane $x/a + y/b + z/c = 1$, as a double integral.

12. A cylindrical column stands on the area bounded by the parabolas $y = x^2$ and $x = y^2$, as base, and is bounded above by $z = 12 - y - x^2$. Find its volume.

13. Find the volume enclosed by the elliptic paraboloid $z = 1 - (x^2 + \frac{1}{2}y^2)$ and the xy -plane.

14. Find the volume cut off from the paraboloid of Example 2 by the plane $z = y$.

15. A cylindrical column stands on the area in the first quadrant bounded by the axes and the hypocycloid, $x^{\frac{2}{3}} + y^{\frac{2}{3}} = 2^{\frac{2}{3}}$, as base, and is bounded above by the hyperbolic paraboloid $z = xy$. Find its volume.

5. Iterated Integrals in Polar Coordinates. The double integral $\iint_S f(x, y) dA$ is by no means bound to a rectangular coordinate system. Indeed, it is by arbitrary choice that we interpret x and y as rectangular

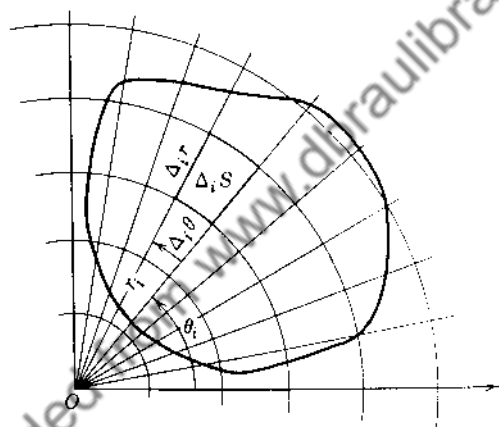


FIG. 13

coordinates. The essential ingredients of this integral are two. First, we have a numerical domain within which the independent variables x and y vary continuously; it is convenient to describe this domain geometrically as a part of a plane. Second, we have a function of x and y that is continuous over the domain. In the formation of the sum $\sum_{i=1}^n f(x_i, y_i) \Delta_i A$, leading to the integral, $\Delta_i A$ denotes the measure of a subdivision $\Delta_i S$ of the domain, and it is required that (x_i, y_i) lie within $\Delta_i S$. The geometrical description of the domain makes it possible to conceive $\Delta_i A$ as an area. The formula by which $\Delta_i A$ is expressed in terms of x and y depends upon the kind of subdivision to which it applies and the geometrical interpretation of these variables. Thus, if x and y are rectangular coordinates, it is natural to subdivide the domain into

rectangular parts with $\Delta_i A = \Delta_i x \cdot \Delta_i y$. In this case we are led to the identification of the double integral with an iterated integral, such as

$$\int_c^d dy \int_{\phi(y)}^{\psi(y)} f(x, y) dx.$$

If x and y were polar coordinates, one would be inclined to subdivide the domain by concentric circles and radial lines, as in Fig. 13. From habit, one would also replace x and y by θ and r . In order to find an expression for $\Delta_i A$, which is, of course not equal to $\Delta_i r \cdot \Delta_i \theta$, we note that $\Delta_i A = \frac{\Delta_i \theta}{2\pi} \cdot R_i$, where R_i denotes the area of the ring with inner radius r_i and outer radius $r_i + \Delta_i r$. Now,

$$\begin{aligned} R_i &= \pi(r_i + \Delta_i r)^2 - \pi r_i^2 = \pi(r_i + \Delta_i r + r_i)(r_i + \Delta_i r - r_i) \\ &= \pi \cdot 2r'_i \Delta_i r, \end{aligned}$$

where r'_i is the average of r_i and $r_i + \Delta_i r$. Hence,

$$\Delta_i A = \frac{\Delta_i \theta}{2\pi} \cdot 2\pi r'_i \Delta_i r = r'_i \Delta_i r \Delta_i \theta.$$

Since, in the sum, $\sum_1^n f(r_i, \theta_i) \Delta_i A$, which leads to the integral $\iint_S f(r, \theta) dA$, we are free to choose r_i and θ_i at will within the subregion $\Delta_i S$, we may put each equal to the average of its least and greatest values in the subregion, viz., r'_i and θ'_i . Hence,

$$\sum_1^n f(r'_i, \theta'_i) \Delta_i A = \sum_1^n f(r'_i, \theta'_i) r'_i \Delta_i r \Delta_i \theta.$$

The limit of the sum on the left is $\iint_S f(r, \theta) dA$, and that of the sum on the right is either of the iterated integrals, $\int_a^b dr \int_{\phi(r)}^{\psi(r)} f(r, \theta) r d\theta$ or $\int_\alpha^\beta d\theta \int_{g(\theta)}^{h(\theta)} f(r, \theta) r dr$, where the limits are obtained from the polar expressions for the boundaries of S . Thus,

$$\iint_S f(r, \theta) dA = \int_a^b dr \int_{\phi(r)}^{\psi(r)} f(r, \theta) r d\theta = \int_\alpha^\beta d\theta \int_{g(\theta)}^{h(\theta)} f(r, \theta) r dr. \quad (5.1)$$

Formula 5.1 furnishes a means of transforming a double integral from rectangular coordinates to polar coordinates. Thus

$$\iint_S F(x, y) dA = \iint_S F(r \cos \theta, r \sin \theta) dA.$$

Now,

$$\iint_S F(x, y) dA = \iint_S F(x, y) dx dy,$$

considered as an iterated integral in either order. Also,

$$\iint_S F(r \cos \theta, r \sin \theta) dA = \iint_S F(r \cos \theta, r \sin \theta) r dr d\theta,$$

considered as an iterated integral in either order. Therefore,

$$\iint_S F(x, y) dx dy = \iint_S F(r \cos \theta, r \sin \theta) r dr d\theta. \quad (5.2)$$

One must not forget to include the factor r in the integrand, in addition to replacing x and y by $r \cos \theta$ and $r \sin \theta$, respectively.

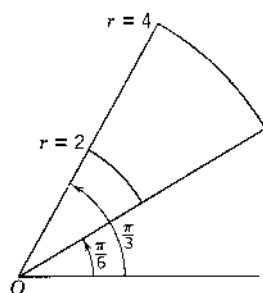


FIG. 14

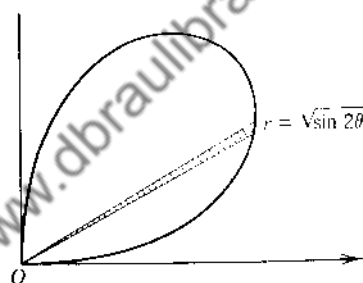


FIG. 15

EXAMPLE 1. Evaluate $\iint_S r^2 \sin \theta dA$, if S is the domain in the first quadrant bounded by the circles $r = 2$ and $r = 4$ and the radial lines $\theta = \pi/6$ and $\theta = \pi/3$ (Fig. 14).

SOLUTION. By Eq. 5.1,

$$\begin{aligned} \iint_S r^2 \sin \theta dA &= \int_{\pi/6}^{\pi/3} d\theta \int_2^4 r^3 \sin \theta dr \\ &= \int_{\pi/6}^{\pi/3} 60 \sin \theta d\theta = 30(\sqrt{3} - 1). \end{aligned}$$

We determine the limits by noting that, if θ were fixed at any value in the domain, r could vary from 2 to 4. In the function of θ resulting from the first integration, θ varies from $\pi/6$ to $\pi/3$.

EXAMPLE 2. Evaluate $\iint_S r^2 \sin \theta dA$ if S is the domain in the first quadrant bounded by the lemniscate $r^2 = \sin 2\theta$ (Fig. 15).

SOLUTION. As in Example 1, $\iint_S r^2 \sin \theta \, dA = \iint_S r^3 \sin \theta \, dr \, d\theta$. In order to find the limits, we note that if θ were fixed at any value in the domain, r could vary from 0 to the value which it assumes on the boundary, i.e., $r = \sqrt{\sin 2\theta}$. A first integration with respect to r corresponds to a summing up along such a radial line, the second integration then being the limit of a summation with respect to θ from 0 to $\pi/2$. Hence,

$$\begin{aligned} \iint_S r^3 \sin \theta \, dr \, d\theta &= \int_0^{\pi/2} d\theta \int_0^{\sqrt{\sin 2\theta}} r^3 \sin \theta \, dr \\ &= \int_0^{\pi/2} \frac{1}{4} \sin^2 2\theta \sin \theta \, d\theta \\ &= \frac{2}{15}. \end{aligned}$$

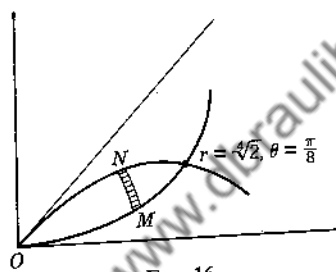


FIG. 16

EXAMPLE 3. A cylindrical solid is based on the area in the first quadrant common to the lemniscates, $r^2 = 2 \sin 2\theta$ and $r^2 = 2 \cos 2\theta$, and is cut off by the surface $z = x^2 - y^2$. Find its volume.

SOLUTION. This volume is $V = \iint_S (x^2 - y^2) \, dA$, where S is the area shown in Fig. 16. The form of the boundary curves suggests that this integral be expressed in polar coordinates. Thus,

$$\begin{aligned} V &= \iint_S (r^2 \cos^2 \theta - r^2 \sin^2 \theta) r \, dr \, d\theta \\ &= \iint_S r^3 \cos 2\theta \, dr \, d\theta. \end{aligned}$$

Integration with respect to r first will require that this integral be broken into two, the upper limit of r in one lying on the curve $r = \sqrt{2 \sin 2\theta}$, and that in the other lying on the curve $r = \sqrt{2 \cos 2\theta}$. This can be avoided by integrating first with respect to θ . Integration with respect to θ , with r temporarily fixed, corresponds to a summing along a circular arc

from one boundary curve to the other, as indicated by the arc MN in Fig. 16. It is clear that if r is fixed at any value in the region, the corresponding circular arc begins in $r^2 = 2 \sin 2\theta$ and ends in $r^2 = 2 \cos 2\theta$, and, hence, a single integration with respect to θ , followed by integration with respect to r , will suffice. The limits for θ are found from the equations of the curves, and are $\frac{1}{2} \sin^{-1}(r^2/2)$ and $\frac{1}{2} \cos^{-1}(r^2/2)$, respectively; since the curves intersect at $\theta = \pi/8$, $r = \sqrt[4]{2}$, the limits for r are 0 and $\sqrt[4]{2}$. Thus,

$$V = \int_0^{\sqrt[4]{2}} dr \int_{\frac{1}{2} \sin^{-1}(r^2/2)}^{\frac{1}{2} \cos^{-1}(r^2/2)} r^3 \cos 2\theta \, d\theta$$

The first integration yields

$$\begin{aligned} & \frac{r^3}{2} \left[\sin 2\theta \right]_{\frac{1}{2} \sin^{-1}(r^2/2)}^{\frac{1}{2} \cos^{-1}(r^2/2)} = \frac{r^3}{2} \left[\sin \left(\cos^{-1} \frac{r^2}{2} \right) - \sin \left(\sin^{-1} \frac{r^2}{2} \right) \right] \\ &= \frac{r^3}{2} \left(\frac{\sqrt{4-r^4}}{2} - \frac{r^2}{2} \right) = \frac{1}{4} [(4-r^4)^{1/2} r^3 - r^5] \end{aligned}$$

Now,

$$\begin{aligned} \int_0^{\sqrt[4]{2}} \frac{1}{4} [(4-r^4)^{1/2} r^3 - r^5] \, dr &= \frac{1}{4} \left[-\frac{1}{6} (4-r^4)^{3/2} - \frac{1}{6} r^6 \right]_0^{\sqrt[4]{2}} \\ &= (2 - \sqrt{2})/6. \end{aligned}$$

EXERCISES

Evaluate the double integrals in Exercises 1-5, in polar coordinates.

- $\iint_S \sin \theta \, dA$; S bounded by the circle $r = 2 \cos \theta$.
- $\iint_S r^2 \, dA$; S the same as in Exercise 1.
- $\iint_S r \sin^2 \theta \, dA$; S the same as in Exercise 1.
- $\iint_S \frac{\cos \theta}{r} \, dA$; S is the right-hand half of the circle $r = 2 \cos \theta$.
- $\iint_S r \cos \theta \, dA$; S is the area outside the circle $r = a$ and inside the circle $r = 2a \cos \theta$.
- Evaluate $\iint_S \frac{x}{\sqrt{x^2 + y^2}} \, dA$, over the right-hand loop of the lemniscate $(x^2 + y^2)^2 = x^2 - y^2$.

7. Evaluate $\iint_S r^2 \cos^2 \theta \, dA$, over the square bounded by the rectangular coordinate axes and the lines $x = 1$, $y = 1$.

8. By setting $A = \iint_S dA$, obtain the formula for area in polar coordinates in Chapter 15.

9. Supply the limits in the iterated integral $\int dr \int f(r, \theta) r \, d\theta$, over the domain common to the circles $r = 2 \cos \theta$ and $r = 2 \sin \theta$.

10. Evaluate $\iint_S \sin 2\theta \, dA$ by means of an iterated integral like that in Exercise 9, over the domain bounded on the right by the cardioid $r = 1 - \cos \theta$ and on the left by the circle $r = \sin \theta$.

11. Evaluate $\iint_S e^{x^2+y^2} \, dA$, if S is the interior of the circle $x^2 + y^2 = 1$.

12. Evaluate $\iint_S e^{-x^2-y^2} \, dA$ over the interior of the circle $x^2 + y^2 = a^2$, and show that the result tends to π as a tends to ∞ .

6. The Area of a Surface. We found that the length of a curve could be calculated by a single integration. Similarly, the area of a curved surface can be found by means of a double integral. Suppose a cylindrical

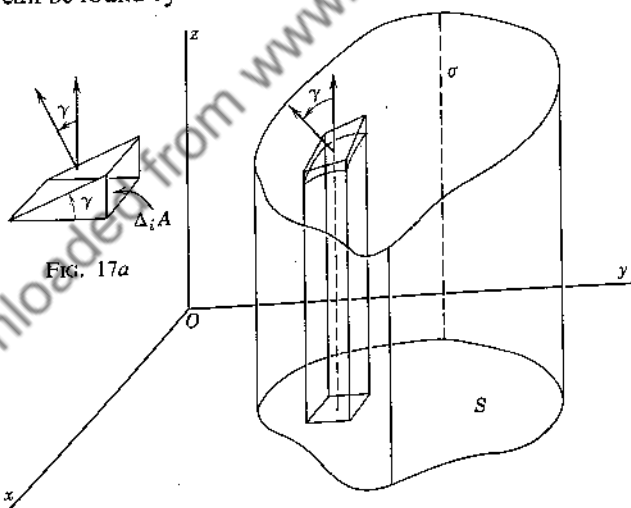


FIG. 17

surface is erected on the boundary of a domain S of the xy -plane, cutting out a segment of the surface $z = f(x, y)$ (Fig. 17). We require the area, σ (sigma), of this segment.

First, imagine a small subdivision of S , $\Delta_i S$, with area $\Delta_i A$, in this subdivision a point (x_i, y_i) with the corresponding height of the surface z_i , and at the point (x_i, y_i, z_i) of the surface, the tangent plane. The cylindrical column based on $\Delta_i S$ cuts out a section of the given surface and a corresponding section of the tangent plane. The area of the section of the tangent plane is $\Delta_i A \cdot \sec \gamma_i$, where γ_i is the angle between the normal at (x_i, y_i, z_i) and the z -axis (Fig. 17a).

Now, if S were completely covered by a network of n subdivisions, and the surface were approximated over each by a section of a tangent plane, as for the section illustrated, the sum of the areas of the tangent sections thus formed would approximate our intuitive conception of the area of the surface itself. Hence, we define this area by the equation

$$\sigma = \lim_{n \rightarrow \infty} \sum_{i=1}^n \Delta_i A \sec \gamma_i = \iint_S \sec \gamma \, dA. \quad (6.1)$$

In order to evaluate this integral, it is necessary to find an expression for $\sec \gamma$ from the equation of the surface, $z = f(x, y)$. We learned, in Chapter 24, that the direction cosines of the normal to the surface at a point (x_i, y_i, z_i) were proportional to $f_x(x_i, y_i)$, $f_y(x_i, y_i)$, -1 . It follows that $\cos \gamma = 1/\sqrt{f_x^2 + f_y^2 + 1}$, and

$$\sec \gamma = \sqrt{f_x^2 + f_y^2 + 1}. \quad (6.2)$$

In this formula, the choice of the positive sign of the square root corresponds to the assumption that γ is acute, and this, in turn, is associated with the idea of the upper side of the surface as positive.

If the equation of the surface is given in implicit form, $F(x, y, z) = 0$, we found in Art. 10, Chapter 24, that the direction cosines of the normal were proportional to F_x , F_y , F_z . Hence,

$$\sec \gamma = \left| \frac{\sqrt{F_x^2 + F_y^2 + F_z^2}}{F_z} \right|. \quad (6.3)$$

It is sometimes convenient to project a surface on the xz -plane or the yz -plane, instead of on the xy -plane as above. For a segment of the surface $F(x, y, z)$ projected on a domain S of the xz -plane, for example, the area is given by

$$\sigma = \iint_S \sec \beta \, dA,$$

where

$$\sec \beta = \left| \frac{\sqrt{F_x^2 + F_y^2 + F_z^2}}{F_y} \right|. \quad (6.4)$$

EXAMPLE 1. Find the area of the conical surface, $x^2 + y^2 - 4z^2 = 0$, within the cylinder $x^2 + y^2 - 4x = 0$, and lying above the xy -plane.

SOLUTION. The domain S is the circle $(x-2)^2 + y^2 = 4$ in the xy -plane (see Fig. 18), and the equation of the surface, a part of whose area is to be calculated, is

$$F(x, y, z) = x^2 + y^2 - 4z^2 = 0. \quad (6.5)$$

Hence,

$$F_x = 2x, \quad F_y = 2y, \quad \text{and} \quad F_z = -8z,$$

and, from (6.3)

$$\sec \gamma = \frac{1}{8z} \sqrt{4x^2 + 4y^2 + 64z^2} = \frac{1}{4z} \sqrt{x^2 + y^2 + 16z^2}.$$

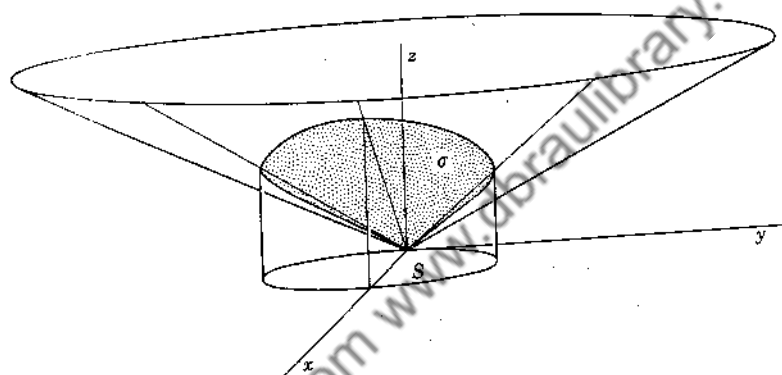


FIG. 18

The required area is then

$$\sigma = \iint_S \frac{1}{4z} \sqrt{x^2 + y^2 + 16z^2} dA,$$

in which z has the value $z = \frac{1}{2} \sqrt{x^2 + y^2}$, from (6.5). Therefore,

$$\begin{aligned} \sigma &= \iint_S \frac{\sqrt{5x^2 + 5y^2}}{2\sqrt{x^2 + y^2}} dA = \iint_S \frac{\sqrt{5}}{2} dA \\ &= \frac{\sqrt{5}}{2} \int_0^4 dx \int_{-\sqrt{4-(x-2)^2}}^{\sqrt{4-(x-2)^2}} dy = \sqrt{5} \int_0^4 dx \int_0^{\sqrt{4-(x-2)^2}} dy \\ &= \sqrt{5} \int_0^4 \sqrt{4-(x-2)^2} dx = 2\sqrt{5}\pi. \end{aligned}$$

EXERCISES

1. Find the area of the part of the plane $x + y + z = 4$ intercepted by the parabolic cylinder $4x = y^2$ and the plane $x = 1$.
2. Find the area of the surface of the circular cylinder $y^2 + z^2 = a^2$ that lies between the parabolic cylinder $ax = y^2$ and the yz -plane.
3. Find the area of one nappe of the cone $x^2 + y^2 + z^2 = 0$ intercepted by the elliptic cylinder $x^2 + y^2/4 = 1$.
4. Find the area of the sphere $x^2 + y^2 + z^2 = a^2$ by means of a double integral.
5. Find the area of the cylinder $x^2 + z^2 = a^2$ that is intercepted by the cylinder $y^2 + z^2 = a^2$.
6. Find the area cut out of the hyperbolic paraboloid $z = xy$ by the cylinder $x^2 + y^2 = a^2$. [Use polar coordinates.]
7. Find the area cut out of the paraboloid $z = x^2 + y^2$ by the cylinder $x^2 + y^2 = a^2$.
8. Find the area cut out of the hyperbolic paraboloid $z = xy$ by the cylindrical surface based on one loop of the lemniscate $r^2 = \cos 2\theta$.
9. Find the area of the cylinder $x^2 + y^2 = 2ax$ that is intercepted by the sphere $x^2 + y^2 + z^2 = 4a^2$. [Hint: Project $\frac{1}{2}$ of the area on the xy -plane.]
10. In Exercise 9, find the area of the sphere intercepted by the cylinder.

7. Physical Applications of Double Integrals. Some of the physical applications taken up in Chapters 9 and 16 for special cases, are, in general, double integration problems. These include moment of area, center of gravity and moment of inertia of a lamina, fluid pressure, and force of attraction of a lamina. The underlying ideas discussed there for figures of special shapes can readily be extended to areas with more general boundaries. This will be illustrated in the following examples.

EXAMPLE 1. Find the mass of a circular lamina of radius a , if the density varies directly as the distance from a fixed point on the boundary.

SOLUTION. Although this type of problem was not discussed earlier, it is a natural application of the double integral. Let the equation of the circle be $(x - a)^2 + y^2 = a^2$, and let the fixed point of the boundary be the origin. Then, the density $= k\sqrt{x^2 + y^2}$. Hence, the mass Δm of a small area ΔA is approximated by $k\sqrt{x^2 + y^2} \Delta A$. Therefore,

$$m = \iint_S k\sqrt{x^2 + y^2} dA,$$

the domain S being the interior of the circle. Using the iterated integral in rectangular coordinates, we have

$$m = 2 \int_0^{2a} dx \int_0^{\sqrt{a^2 - (x-a)^2}} k\sqrt{x^2 + y^2} dy.$$

One might expect that the evaluation of this integral would be laborious and that a simplification would result from the use of polar coordinates.

In this formulation, the equation of the circle is $r = 2a \cos \theta$, and

$$\begin{aligned} m &= \iint_S kr \, dA = 2k \int_0^{\pi/2} d\theta \int_0^{2a \cos \theta} r^2 \, dr \\ &= \frac{16a^3 k}{3} \int_0^{\pi/2} \cos^3 \theta \, d\theta = \frac{32}{9} ka^3. \end{aligned}$$

EXAMPLE 2. Find the moment of mass in Example 1, about the y -axis; find the center of gravity of the mass.

SOLUTION. The moment of mass about the y -axis is

$$\begin{aligned} M_y &= \iint_S kx\sqrt{x^2+y^2} \, dA = \iint_S kr \cos \theta \cdot r \, dA \\ &= 2k \int_0^{\pi/2} d\theta \int_0^{2a \cos \theta} r^3 \cos \theta \, dr = 8a^4 k \int_0^{\pi/2} \cos^5 \theta \, d\theta \\ &= \frac{64}{15} a^4 k. \end{aligned}$$

On account of symmetry, the center of gravity lies on the x -axis. Its distance from the origin is

$$\bar{x} = \frac{M_y}{m} = \left(\frac{64}{15} a^4 k \right) / \left(\frac{32}{9} a^3 k \right) = 6a/5.$$

CAUTION. The employment of symmetry about a line, to simplify the calculation of a double integral, as in the above examples, requires that the domain of integration be symmetrical about the line, and, also, that the integrand have the same value at symmetrical points. For example, the domain in Example 1 is symmetric about the x -axis, and the integrand has the same value at any two symmetric points (x, y) and $(x, -y)$.

EXAMPLE 3. A right triangular lamina bounded by the y -axis and the lines $y = x$ and $y = a$ has its density proportional to the distance from the hypotenuse. Find the moment of inertia (a) about the y -axis; (b) about the x -axis; (c) about the origin.

SOLUTION. The distance from the hypotenuse $y = x$ is $(y - x)/\sqrt{2}$, by analytic geometry. Hence (see Chapter 16, Arts. 8 and 9),

$$\begin{aligned} I_y &= \iint_S kx^2 \frac{y-x}{\sqrt{2}} \, dA = \frac{k}{\sqrt{2}} \int_0^a dy \int_0^y x^2 (y-x) \, dx \\ &= \frac{k}{12\sqrt{2}} \int_0^a y^4 \, dy = \frac{ka^5\sqrt{2}}{120}. \end{aligned}$$

$$\begin{aligned}
 I_x &= \iint_S k y^2 \frac{y+x}{\sqrt{2}} dA = \frac{k}{\sqrt{2}} \int_0^a dy \int_0^y y^2 (y+x) dx \\
 &= \frac{k}{2\sqrt{2}} \int_0^a y^4 dy = \frac{ka^5\sqrt{2}}{20}.
 \end{aligned}$$

Since

$$I_0 = \iint_S k(x^2 + y^2) \frac{y}{\sqrt{2}} dA,$$

it follows that

$$I_0 = I_x + I_y = \frac{7ka^5\sqrt{2}}{120}.$$

EXERCISES

Find the mass and center of gravity of each of the laminae in Exercises 1-5.

1. A square bounded by the coordinate axes and the lines $x = b$, $y = b$, if the density is proportional to the square of the distance from 0.
2. A right triangle bounded by the coordinate axes and the line $x + y = b$, if the density is proportional to the square of the distance from 0.

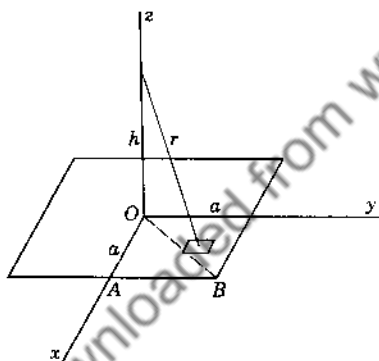


FIG. 19

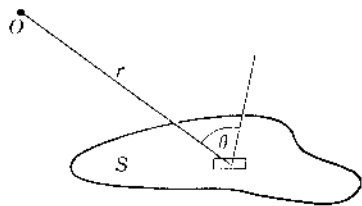


FIG. 20

3. The triangle in Exercise 2, if the density is proportional to the distance from the base.
4. The triangle in Exercise 2, if the density is proportional to the distance from the hypotenuse.
5. The circle $r = 2a \cos \theta$, if the density is inversely proportional to the distance from 0.
6. Find I_x , I_y , and I_0 for the lamina in Exercise 1.
7. Find the moment of inertia of the lamina in Exercise 4, about each side. [Hint: To find I about the hypotenuse, rotate the triangle so that this line lies on the y -axis.]

8. Find the total force of gravitational attraction due to a semicircular lamina of radius 1, and uniform density δ , upon a unit mass located at the intersection of the extended circumference and the axis of symmetry (see Chapter 16, Art. 5).

$$\left[\text{Hint: } F = k \int_S \frac{\cos \theta}{r^2} dA. \right]$$

9. A lamina bounded by the lines $y = 1 + x$, $y = 1 - x$, and $y = 2$ has its density inversely proportional to the distance from the origin. Find the force of attraction which it exerts on a unit mass at the origin.

10. Find the total force of gravitational attraction due to a uniform square lamina, $2a$ units on each side, upon a unit mass located on a line perpendicular to the lamina at its center and h units away (Fig. 19). [Hint: Use polar coordinates and integrate over triangle OAB .]

11. The intensity of light from a point source is inversely proportional to the square of the distance from the source. Show that the total illumination on a surface S due to a source at O (Fig. 20), is equal to the double integral $\iint_S \frac{k}{r^2} \cos \theta dA$, in which r is the variable distance from O to S , and θ is the angle of incidence of the rays. Hence, calculate the total illumination on a circular table top of radius a cast by a point source located h units above its center.

8. An Integral over a Three-Dimensional Domain. A function of three variables, $F(x, y, z)$, that is continuous for all points (x, y, z) in a three-dimensional domain S , has an integral over this domain, defined as follows. We break up the domain into n subdivisions, having the volumes $\Delta_i V$; choose a point (x_i, y_i, z_i) in each subdivision; and form the sum $\sum_{i=1}^n F(x_i, y_i, z_i) \Delta_i V$. We now let n tend to infinity and require that the greatest diameter of each subdivision tend to 0 at the same time. This limit is called the *domain* integral, or *triple* integral, of the function $F(x, y, z)$ over the three-dimensional domain. We write

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n F(x_i, y_i, z_i) \Delta_i V = \iiint_S F(x, y, z) dV.$$

The proof that this limit exists and is the same for all methods of subdivision, subject to the provision that the greatest diameter of each tend to 0 as n tends to infinity, and for all choices of (x_i, y_i, z_i) in $\Delta_i V$, is much like that for double integrals, and will not be given.

There is no geometrical interpretation of a triple integral analogous to the interpretations of single and double integrals as areas and volumes, respectively. Physical interpretations are not hard to find. Perhaps the simplest is the conception of the mass of a body of variable density as a triple integral. Suppose, for example, that we require the mass of a cubical block 4 units on each edge, if the density is equal to $1 + .1r^2$,

where r is the distance from a corner. Placing the block in a rectangular coordinate system, we express the mass by

$$m = \iiint_S [1 + .1(x^2 + y^2 + z^2)] dV,$$

in which S represents the cube.

9. A Three-Dimensional Iterated Integral. Suppose that the function $f(x, y, z)$ is defined and continuous in the domain bounded by the planes

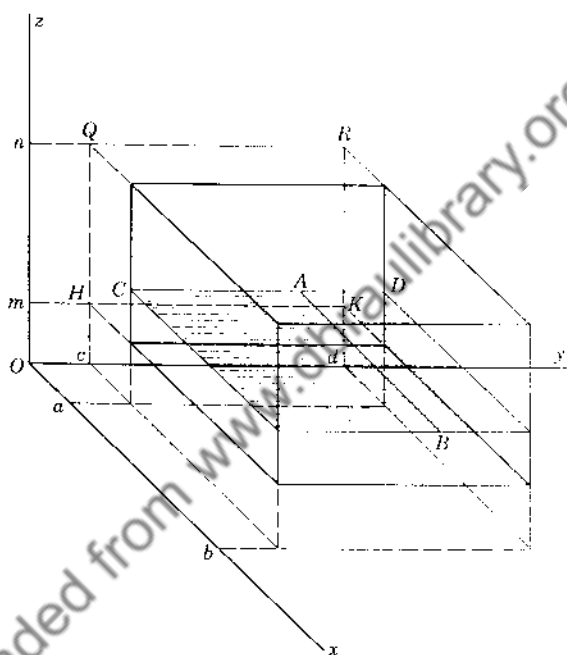


FIG. 21

$x = a$, $x = b$, $y = c$, $y = d$, $z = m$, and $z = n$ (Fig. 21). If the function is integrated with respect to x between limits a and b , with y and z held constant while the integration is performed, the result is a function of y and z . That is,

$$\int_a^b f(x, y, z) dx = g(y, z). \quad (9.1)$$

This integral is the limit of a summation along a line like AB , of products $f(x_i, y_i, z_i) \Delta x$, just as the integral $\int_a^b f(x) dx$ is the limit of a summation along a segment of the x -axis of products $f(x_i) \Delta x$. The result is general, being a formula for the value of such an integral for any line parallel to the x -axis within the domain, i.e., for any choice of y and z in the

rectangular area of the yz -plane, $HKRQ$, $c \leq y \leq d$, $m \leq z \leq n$. This is what is meant by putting it equal to $g(y, z)$.

The function $g(y, z)$ can now be integrated with respect to y , holding z constant. This integral is the limit of a summation along a line parallel to the y -axis, like CD , of products $g(y_j, z_j) \Delta y_j$. Since the value of $g(y_j, z_j)$ used in this summation is already an integral along a line parallel to the x -axis, and the second integration arises from summing up such integrals, this integration may be pictured by the formation of a layer parallel to the xy -plane. It is, in fact, the double integral over such a layer. Since z is fixed in this integration and can be taken arbitrarily as any value of z in the domain, the result is a function of z , giving the value of the double integral over any layer. That is,

$$\int_c^d \left[\int_a^b f(x, y, z) dx \right] dy = \int_c^d g(y, z) dy = h(z). \quad (9.2)$$

We may now integrate $h(z)$,

$$\int_m^n \left\{ \int_c^d \left[\int_a^b f(x, y, z) dx \right] dy \right\} dz = \int_m^n h(z) dz. \quad (9.3)$$

This integral is the limit of a summation along a line parallel to the z -axis of products $h(z_k) \Delta z_k$, i.e., the limit of a summation of values of integrals over all layers.

The triple iterated integral in (9.3) is usually denoted more simply by

$$\int_m^n dz \int_c^d dy \int_a^b f(x, y, z) dx, \quad \text{or} \quad \int_m^n \int_c^d \int_a^b f(x, y, z) dx dy dz,$$

and may be evaluated in any order corresponding to the permutations of dx , dy , and dz .

We now state without proof that a three-dimensional domain integral, as described in Art. 8, can be evaluated by a corresponding triple iterated integral. For example, the integral for the mass in Art. 8 is

$$\begin{aligned} m &= \iiint_S [1 + .1(x^2 + y^2 + z^2)] dV \\ &= \int_0^4 dz \int_0^4 dy \int_0^4 [1 + .1(x^2 + y^2 + z^2)] dx \\ &= \int_0^4 dz \int_0^4 (6.13 + .4y^2 + .4z^2) dy \\ &= \int_0^4 (24.52 + 8.53 + 1.6z^2) dz \\ &= 166.33. \end{aligned}$$

If the domain S is bounded by curved surfaces instead of planes parallel to the coordinate planes, an iterated integral can still be used. It must be expected that the limits of the first two integrations will be variable functions of the remaining variables, obtained from the equations of the bounding surfaces. This is illustrated in the following examples.

EXAMPLE 1. Evaluate $\iiint_S x \, dA$ if S is bounded by the xy -plane, the cylinder $x^2 + y^2 = 4$, and the plane $x/4 + y/6 + z/8 = 1$.

SOLUTION. We put this integral equal to the iterated integral $\int dx \int dy \int x dz$, in which the limits are to be found from the bounding surfaces. We take z as the first variable of integration, since all line segments parallel to the z -axis terminate in a single surface at both ends, viz., $z = 0$ below and $z = 8 - 2x - 4y/3$ above. The line AB in Fig. 22 typifies this integration. The second integration, with respect to y , extends from one side of the cylinder to the other, i.e., from $y = -\sqrt{4-x^2}$ to $y = \sqrt{4-x^2}$. The shaded area suggests this integration. The final integration, with respect to x , extends between the extreme values of x on the cylinder, i.e., from $x = -2$ to $x = 2$. Inserting the limits and computing,

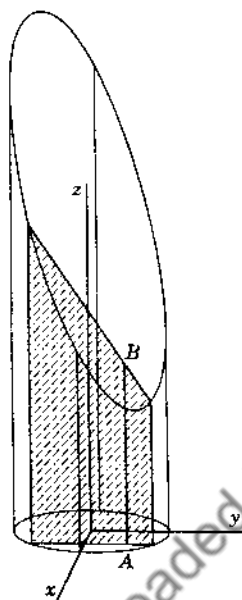


FIG. 22

$$\begin{aligned} \iiint_S x \, dA &= \int_{-2}^2 dx \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} dy \int_0^{8-2x-4y/3} x \, dz \\ &= \int_{-2}^2 dx \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} (8x - 2x^2 - \frac{4}{3}xy) \, dy \\ &= \int_{-2}^2 (16x - 4x^2)\sqrt{4-x^2} \, dx = -8\pi. \end{aligned}$$

EXAMPLE 2. Set up the triple integral $\iiint_S xyz \, dA$ as an iterated integral, if S is the region in the first octant bounded by the paraboloids $4z = x^2 + y^2$ and $4y = x^2 + z^2$ and the yz -plane (Fig. 23).

SOLUTION. Again we integrate first with respect to z between the lower surface $z = (x^2 + y^2)/4$ and the upper surface $z = \sqrt{4y - x^2}$. In order to obtain the limits for x and y we must determine the curve in the xy -plane into which the intersection of the paraboloids projects. On the intersection, both equations hold. By subtraction, we find that $(z - y)(z + y + 4) = 0$. We conclude that $z = y$, since $z + y + 4 \neq 0$, as neither y nor z can be negative. Substituting $z = y$ in the equation

of either paraboloid, in order to obtain the relation between x and y on the intersection, we find that $x^2 + (y - 2)^2 = 4$. The region of the xy -plane in which x and y vary is therefore bounded by the semicircle

$$\begin{aligned} x &= + \sqrt{4 - (y - 2)^2} \\ &= + \sqrt{4y - y^2}, \end{aligned}$$

and the y -axis. We can now write the iterated integral,

$$\iiint_S xyz \, dA = \int_0^4 dy \int_0^{\sqrt{4y-y^2}} dx \int_{(x^2+y^2)/4}^{\sqrt{4y-x^2}} xyz \, dz.$$

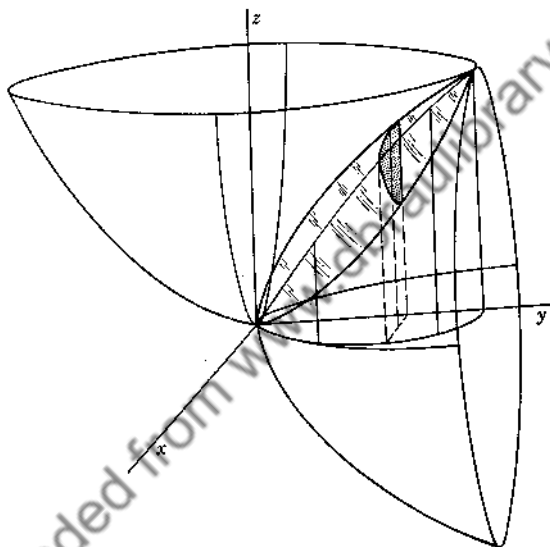


FIG. 23

EXERCISES

1. Calculate the iterated integral in Example 2.
2. Express the volume of a 3-dimensional domain by an integral over that domain, and hence find the volume of the domain in Example 1.
3. Show that the expression for the volume of a three-dimensional domain by a triple integral is immediately reducible to a double integral; and, hence, that there is no advantage in the use of a triple integral.
4. Find the volume of the domain bounded by the two paraboloids $4z = x^2 + y^2$ and $4z + x^2 + y^2 = 2$.

5. Show that $\iiint_S (x^2 + y^2 + z^2)xyz \, dV = 0$, where S is the interior of the sphere $x^2 + y^2 + z^2 = a^2$.

6. Calculate $\iiint_S z \, dV$, where S is the domain bounded above by the sphere $x^2 + y^2 + z^2 = 1$ and below by the cone $x^2 + y^2 = z^2$.

7. Show that $\iiint_S f(x, y)z^{2n-1} \, dV$ is equal to 0 if S is symmetric with respect to the xy -plane. Hence, show that the integral in Exercise 6 is equal to 0 if the domain S is the whole solid inside both the cone and the sphere. What can we conclude about this integral over the part of this domain for which $z < 0$?

8. Calculate $\iiint_S (x + y + z) \, dV$, where S is bounded by the plane $x + y + z = 1$ and the coordinate planes.

9. Show that

$$\int_{x_0}^{x_1} dx \int_{y_0}^{y_1} dy \int_{z_0}^{z_1} f(x)g(y)h(z) \, dz = \int_{x_0}^{x_1} f(x) \, dx \cdot \int_{y_0}^{y_1} g(y) \, dy \cdot \int_{z_0}^{z_1} h(z) \, dz.$$

Apply this result to find the mass of the block in Art. 8, if the density varies as e^{x+y+z} .

10. Find the mass of the spherical solid bounded by $x^2 + y^2 + (z - a)^2 = a^2$, if the density varies directly as $|xyz|$. [Hint: Find the mass of the octant where x, y , and z are all positive.]

10. Cylindrical and Spherical Coordinates. Since cylindrical coordinates are a combination of polar coordinates in a plane, with a third rectangular coordinate attached, the iterated triple integral in cylindrical coordinates for a given three-dimensional domain integral takes a form analogous to the iterated double integral in polar coordinates. The element of volume ΔV is simply the product of the element of plane area $r \, \Delta \theta \, \Delta r$, multiplied by Δz (see Fig. 24). That is,

$$\iiint_S f(r, \theta, z) \, dV = \iiint_S f(r, \theta, z) r \, d\theta \, dr \, dz,$$

with appropriate limits of integration, in any order. From this result, we have a formula for change of variables in a triple integral from rectangular coordinates to cylindrical coordinates,

$$\iiint_S F(x, y, z) \, dx \, dy \, dz = \iiint_S f(r, \theta, z) r \, d\theta \, dr \, dz, \quad (10.1)$$

in which $f(r, \theta, z) = F(r \cos \theta, r \sin \theta, z)$.

Frequently, the form of the integrand makes a change to cylindrical coordinates appear advantageous. For example,

$$\begin{aligned} \iiint_S (a^2 - x^2 - y^2)^{\frac{1}{2}} z \tan^{-1}(y/x) \, dV &= \iiint_S (a^2 - r^2)^{\frac{1}{2}} z \, \theta \, dr \, dz \\ &= \iiint_S (a^2 - r^2)^{\frac{1}{2}} z \, \theta \, r \, dr \, d\theta \, dz. \end{aligned}$$

The use of cylindrical coordinates may also be suggested by the boundary surfaces. Thus, $\iiint_S F(x, y, z) dV$, in which S is a circular cylindrical solid with axis on the z -axis and radius a , becomes

$$\int_{z_0}^{z_1} dz \int_0^{2\pi} d\theta \int_0^a F(r \cos \theta, r \sin \theta, z) r dr.$$

Although cylindrical coordinates are suggested in these examples, their use is not necessarily advantageous. In the first example the boundaries

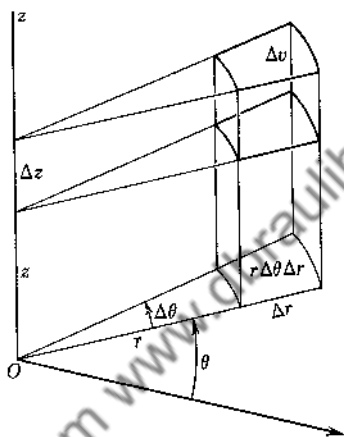


FIG. 24

of S might lead to unmanageable limits, and in the second example, $F(r \cos \theta, r \sin \theta, z)$ might be so complicated a function of r , θ , and z as to nullify the advantage gained by the simple limits.

EXAMPLE 1. The density of a circular cylinder varies as the distance from the axis. Find its mass.

SOLUTION. Letting the radius be a and the length h , the mass is

$$\begin{aligned} m &= \iiint_S kr dV = \int_0^h dz \int_0^{2\pi} d\theta \int_0^a kr^2 dr \\ &= 2\pi ka^3 h / 3. \end{aligned}$$

EXAMPLE 2. Evaluate $\iiint_S (a^2 - x^2 - y^2)^{\frac{1}{2}} z \tan^{-1} \frac{y}{x} dV$ if S is the interior of the hemisphere $z = \sqrt{a^2 - x^2 - y^2}$.

SOLUTION. Putting this integral in cylindrical coordinates, it becomes

$$\iiint_S (a^2 - r^2)^{\frac{1}{2}} z \, \theta \, dV = \int d\theta \int dz \int (a^2 - r^2)^{\frac{1}{2}} r \, z \, dr.$$

In order to find the limits, we imagine the integration with respect to r as taking place along a radial line from the z -axis to the spherical boundary (Fig. 25); i.e., from $r = 0$ to $r = \sqrt{a^2 - z^2}$. Integration with respect to z then corresponds to summing up such results in a layer from $z = 0$ to

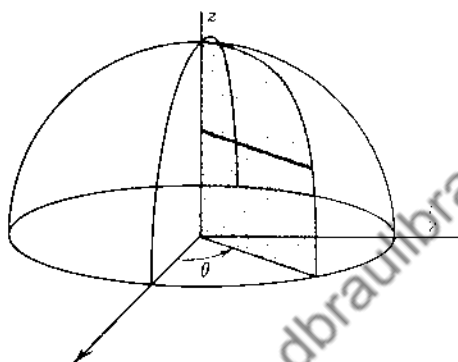


FIG. 25

$z = a$. Finally, integration with respect to θ sums up results corresponding to all layers radiating from the z -axis. Hence, the iterated integral is

$$\begin{aligned} \int_0^{2\pi} d\theta \int_0^a dz \int_0^{\sqrt{a^2 - z^2}} (a^2 - r^2)^{\frac{1}{2}} \theta z r \, dr &= \int_0^{2\pi} d\theta \int_0^a \frac{a^5 - z^5}{5} \theta z \, dz \\ &= \int_0^{2\pi} \frac{5a^7}{14} \theta \, d\theta = \frac{5\pi^2 a^7}{7}. \end{aligned}$$

Spherical coordinates, defined in Art. 14, Chapter 23, suggest a subdivision of a three-dimensional domain by spheres centered at the origin, cones having the z -axis as a common axis, and half-planes through the z -axis. A typical element of such a subdivision is shown in Fig. 26. The coordinates of P are (r, θ, ϕ) , and the lengths of the edges of the inner bounding surface are $PQ = r\Delta\phi$, $PR = r\sin\phi \cdot \Delta\theta$, and the thickness is $PT = \Delta r$. Thus, one would expect ΔV to be approximately equal to $r^2 \sin\phi \Delta r \Delta\phi \Delta\theta$. In fact, although we shall not prove it here, $\Delta V = (r')^2 \sin\phi' \Delta r \Delta\phi \Delta\theta$, in which r' and ϕ' correspond to a point within ΔV . Consequently, the triple integral $\iiint_S f(r, \theta, \phi) \, dV$ is equal to the iterated integral $\int \int \int_S f(r, \theta, \phi) r^2 \sin\phi \, dr \, d\phi \, d\theta$, in any of the six

possible orders of integration, with limits determined from the bounding surfaces of S .

As a formula for changing a triple integral from rectangular coordinates to spherical coordinates, we have

$$\iiint_S F(x, y, z) \, dx \, dy \, dz = \iiint_S G(r, \theta, \phi) r^2 \sin \phi \, dr \, d\phi \, d\theta,$$

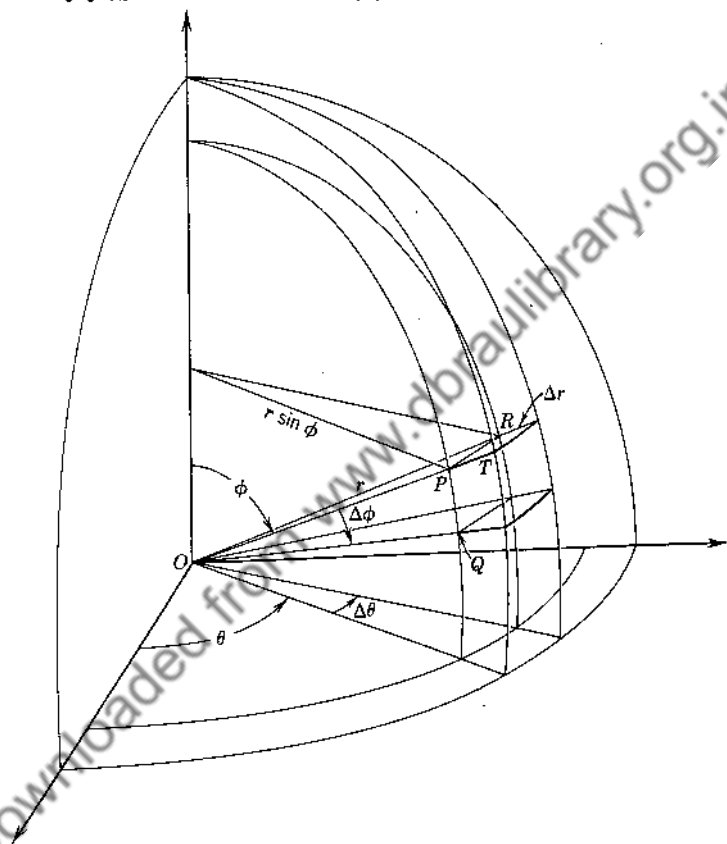


FIG. 26

in which $G(r, \theta, \phi) = F(r \sin \phi \cos \theta, r \sin \phi \sin \theta, r \cos \phi)$ (see Art. 14, Chapter 23).

The use of spherical coordinates may be suggested, as with cylindrical coordinates, by the nature of the bounding surfaces or by the form of the integrand.

EXAMPLE 3. Find the mass of a sphere of radius a if the density varies as the distance from the center.

SOLUTION. We have

$$\begin{aligned} m &= \iiint_S \delta r \, dV \\ &= \delta \int_0^{2\pi} d\theta \int_0^{\pi} d\phi \int_0^a r^3 \sin \phi \, dr \\ &= \pi a^4. \end{aligned}$$

Note that the vertex angle of the cone, i.e., ϕ , varies from 0 to π , whereas θ varies from 0 to 2π .

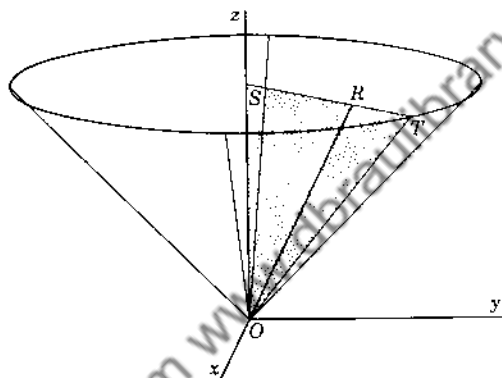


FIG. 27

EXAMPLE 4. Evaluate $\iiint_S \frac{\cos \phi}{r^2} \, dV$, if S is the conical solid bounded by the surface $x^2 + y^2 = z^2$, and the plane $z = h$ (Fig. 27).

SOLUTION. In order to obtain the limits for the iterated integral, we note that r varies from 0 to the plane $z = h$, i.e., from 0 to $h \sec \phi$, along a line OR ; ϕ varies from 0 to the vertex angle of the cone, which is $\pi/4$ (found from the equation of the cone), corresponding to the layer OST ; θ varies from 0 to 2π . Hence,

$$\begin{aligned} \iiint_S \frac{\cos \phi}{r^2} \, dV &= \int_0^{2\pi} d\theta \int_0^{\pi/4} d\phi \int_0^{h \sec \phi} \sin \phi \cos \phi \, dr \\ &= \int_0^{2\pi} d\theta \int_0^{\pi/4} h \sin \phi \, d\phi \\ &= \int_0^{2\pi} h(1 - \tfrac{1}{2}\sqrt{2}) \, d\theta = \pi h(2 - \sqrt{2}). \end{aligned}$$

11. Applications of Triple Integrals. Among the applications of triple integrals are some that were handled by means of single integrals in Chapter 16, due to the special nature of the domain. The determination of total mass, moment of mass, center of gravity, moment of inertia, and force of attraction are not essentially different in three dimensions from the same problems in one or two dimensions. Some examples will bring this out.

EXAMPLE 1. Find the center of gravity of the pyramidal solid bounded by the plane $x + y + z = 1$ and the coordinate planes, if the density at a point varies directly as the sum of the coordinates.

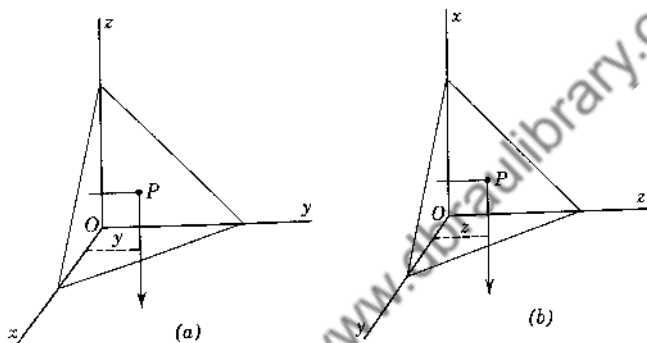


FIG. 28

SOLUTION. It follows immediately from the result of Exercise 8, Art. 9, that the mass of this solid is $k/8$. For its moment with respect to the xz -plane, we have

$$M_{xz} = \iiint_S k(x + y + z)y \, dV.$$

This moment can be conceived physically as the moment the solid would exert about any horizontal axis in the xz -plane, in particular, the x -axis (Fig. 28a). Similarly, the moment with respect to the yz -plane is

$$M_{yz} = \iiint_S k(x + y + z)x \, dV.$$

For the moment with respect to the xy -plane we imagine this plane placed vertically, so that the weight of the solid would exert a moment about any horizontal axis in the xy -plane, e.g., the y -axis in Fig. 28b. Therefore,

$$M_{xy} = \iiint_S k(x + y + z)z \, dV.$$

Evaluating these integrals, we obtain

$$M_{xz} = M_{yz} = M_{xy} = k/30.$$

Therefore, the coordinates of the center of gravity are

$$\bar{x} = (M_{yz})/m = (k/30)/(k/8) = 4/15, \quad \bar{y} = \dots$$

EXAMPLE 2. Find the moment of inertia of a homogeneous tetrahedron bounded by the coordinate planes and the plane $x/a + y/b + z/c = 1$, about each of the three axes.

SOLUTION. For the moment of inertia about the z -axis we have

$$\begin{aligned} I_z &= k \iiint_S (x^2 + y^2) dV \\ &= k \int_0^a dx \int_0^{b(1-x/a)} dy \int_0^{c(1-y/b-x/a)} (x^2 + y^2) dz \\ &= (kabc/60)(a^2 + b^2). \end{aligned}$$

Similarly,

$$I_y = (kabc/60)(a^2 + c^2),$$

and

$$I_x = (kabc/60)(b^2 + c^2).$$

A concept closely related to gravitational attraction is that of potential. Imagine a body of unit mass at the origin O , exerting a force of attraction upon any mass-point (of mass m) in a domain of space. At various points in this domain the forces differ both in amount and direction. At a specific point r units from O , it is km/r^2 in the direction toward O . We now imagine a function over this domain, $u(x, y, z)$, having its derivative in the direction, α , toward O equal to km/r^2 , i.e., $D_{(\alpha)}u(x, y, z) = km/r^2$. This function $u(x, y, z)$, is called a *potential function* of the force field under discussion. In particular, for a point on the x -axis, $D_{(\alpha)}u = -D_xu$, the sign being minus because α is the direction toward the origin. Hence

$$D_xu = -km/x^2, \quad \text{and} \quad u = km/x + C.$$

The constant C is put equal to 0 in accordance with the physical assumption that $u = 0$ when x becomes infinite. Hence, $u = km/x$. If we suppose that the units are so chosen that $k = 1$, we have $u = m/x$.

An equivalent definition specifies that the potential u_1 of a mass-point m_1 at x_1 on the x -axis is the work done against the force of attraction between m_1 and a unit mass at the origin in moving m_1 from its position

at x_1 to an infinite distance away. Since the force of attraction is $-m_1/x^2$, the force that does the work is m_1/x^2 , and we have for u_1

$$u_1 = \int_{x_1}^{\infty} (m_1/x^2) dx = \lim_{A \rightarrow \infty} \left(-m_1/x \right)_{x_1}^A = m_1/x_1.$$

This definition is extended to the potential at a point P due to a system of mass-points by the equation,

$$u = m_1/\rho_1 + m_2/\rho_2 + \dots + m_n/\rho_n,$$

where $\rho_1, \rho_2, \dots, \rho_n$ denote the distances of the mass-points $m_1, m_2, \dots, m_n$, respectively, from P . The potential of a continuous distribution of mass through a three-dimensional domain is, then,

$$u = \iiint_S \delta/\rho \, dV,$$

where δ denotes the density at any point and ρ is the variable distance from the fixed point P .

Potential, unlike force, is a scalar quantity. Since it is not a vector, there is no question of components.

EXAMPLE 3. Find the potential of a homogeneous shell bounded by concentric spheres, at an interior point. Let the inner radius be a , and the outer radius be c .

SOLUTION. At the center, which we place at the origin of the coordinate system,

$$\begin{aligned} u &= \iiint_S \delta/r \, dV = \delta \int_0^{2\pi} d\theta \int_0^\pi d\phi \int_a^c (\delta/r) r^2 \sin \phi \, dr \\ &= 2\pi\delta(c^2 - a^2). \end{aligned}$$

Now, let us take P at any point on the z -axis, but at a point inside the shell, say at $z = h < a$ (Fig. 29). By the law of cosines,

$$\rho^2 = r^2 + h^2 - 2rh \cos \phi.$$

Writing the iterated integral for u , with the integration first with respect to ϕ ,

$$u = \delta \int_0^{2\pi} d\theta \int_a^c dr \int_0^\pi \frac{r^2 \sin \phi}{\sqrt{r^2 + h^2 - 2rh \cos \phi}} d\phi.$$

Now,

$$\begin{aligned} \int_0^\pi \frac{r^2 \sin \phi}{\sqrt{r^2 + h^2 - 2rh \cos \phi}} d\phi &= \int_0^\pi \frac{r}{2h} (r^2 + h^2 - 2rh \cos \phi)^{-\frac{1}{2}} 2rh \sin \phi d\phi \\ &= \left[\frac{r}{h} \sqrt{r^2 + h^2 - 2rh \cos \phi} \right]_0^\pi = \frac{r}{h} (r + h - r + h) = 2r. \end{aligned}$$

Therefore,

$$u = \delta \int_0^{2\pi} d\theta \int_a^r 2r \, dr = 2\pi\delta a^2. \quad (3)$$

This result is independent of h and agrees with that obtained for P at the origin. Since, for any location of P inside the shell, the coordinate system could be oriented so that P would lie on the z -axis, it follows that the potential due to the shell is the same at all interior points.

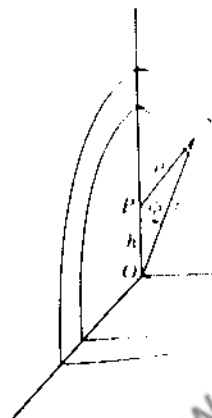


FIG. 29

EXERCISES

1. Carry out the integration in Example 2.
2. Find the mass of a cone of revolution of altitude h , radius a , if the density varies as the distance from the axis.
3. Find the mass in Exercise 2 if the density varies as the distance from the vertex.
4. Find the mass of a sphere of radius a , if the density varies as the absolute value of the product of the distances from three mutually perpendicular planes through the center.
5. Find the mass of the ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, if the density varies as the absolute value of the product of the distances from the coordinate planes.
6. Find the center of gravity of the octant of a sphere of radius a , if the density varies as the distance from the vertex.
7. Find the center of gravity in Exercise 6, if the density varies as the distance from one plane face.
8. Find the center of gravity of the cone in Exercise 2.
9. Find the moment of inertia of the cone in Exercise 2 about its axis.
10. Find the moment of inertia of the solid in Exercise 6 about any one of its straight edges.
11. Find the moment of inertia of the cone in Exercise 3 about its axis.
12. Find the total force of attraction of an octant of a sphere of radius a upon a unit mass at its vertex, if the density is uniform.
13. Find the potential of the solid in Exercise 12 at the vertex.
14. Find the total force of attraction exerted by a solid circular cone of constant density upon a unit mass at its vertex.

15. Find the potential of the cone in Exercise 14 at the vertex.
16. Find the total force of attraction of a homogeneous sphere of radius a upon a unit mass on its surface.
17. Find the total force of attraction exerted by a circular cylindrical solid of uniform density, upon a unit mass located on the axis of the cylinder extended. Let the radius $= a$, the altitude $= h$, and the distance of the unit mass from the nearest base $= b$. [Hint: Use cylindrical coordinates. Compare with Exercise 6, Art. 5, Chapter 16.]
18. Find the total force of attraction of a hemispherical solid of radius a on a unit mass at the center of the base, if the density varies as the distance from the base.
19. Show that the potential of the shell in Example 3, at an exterior point, is the same as though the entire mass were concentrated at the center of the shell. Hence, what can be said of the potential of a solid homogeneous sphere at an exterior point? [Use Fig. 29 and place P on the z -axis at $z = b > c$.]
20. Find the total force of attraction of a solid bounded by the paraboloid $4az = x^2 + y^2$ and the plane $z = a$ upon a unit mass at the origin.

12. Summary and Conclusion. Several results have been given in this chapter without proof. Foremost among them are the theorems on the existence of domain integrals and the theorems that equate domain integrals with iterated integrals. Among the applications, also, certain integrals have been assumed to apply, without rigorous proofs having been given. We have taken up the double integral as essentially a volume problem, as the single integral is an area problem. The triple integral is most readily conceived in physical interpretations, such as mass or total force of attraction. There is no real bar to extending integrals to four-, or n -, dimensional domains; but interpretations are rare, and little or no need for such extensions can be felt at this juncture.

Again, at the end of the book, as at the beginning, the student is reminded that the purpose of a first course is to gain insight and some facility in the use of the concepts and methods of calculus. The first requirement of a demonstration intended to meet this purpose is plausibility—it should seem right. A serious student, however, will not be satisfied with mere plausibility. Nor will he be discouraged because his grasp of the concepts is imperfect and his confidence in dealing with applications incomplete. The subject of calculus is an enormous field. A first course can present the concepts in a meaningful way, and can give a partial appreciation of the powerful methods of calculus. Understanding will be deepened, and methods extended and strengthened, by further study especially focused on fundamental theory. There are several advanced calculus books that are admirable for this purpose.

INTEGRAL FORMULAS AND HINTS FOR THEIR USE

The principal formulas for integration appear in Chapter 13, pp. 282-3. A short list of additional formulas follows, together with hints for adapting integrands to these formulas and those in Chapter 13.

Special conditions governing the use of a formula must be carefully noted.

$$1. \int \sqrt{u^2 \pm a^2} du = \frac{1}{2}[u\sqrt{u^2 \pm a^2} \pm a^2 \log(u \pm \sqrt{u^2 \pm a^2})].$$

$$2. \int \sqrt{a^2 - u^2} du = \frac{1}{2}[u\sqrt{a^2 - u^2} + a^2 \sin^{-1} u/a].$$

Forms involving $u^2 \pm a^2$ or $a^2 - x^2$ are reduced to trigonometric integrals by the substitutions of Art. 9, Chapter 13.

For $a^2 - u^2$, put $u = a \sin v$.

For $u^2 + a^2$, put $u = a \tan v$.

For $u^2 - a^2$, put $u = a \sec v$.

Quadratic expressions, $px^2 + qx + r$, can be brought into one of the above forms by completing the square: $px^2 + qx + r = p(u^2 + a^2)$, in which $u = x + q/(2p)$, $a^2 = (q^2 - 4pr)/(4p)$.

$$3. \int \sqrt{\frac{au + b}{cu + k}} du = \int \frac{au + b}{\sqrt{pu^2 + qu + r}} du, \text{ with } p = ac, q = bc + ak, r = bk.$$

$$4. \int \frac{du}{(u^2 + b)^m} = \frac{u}{2(m-1)(u^2 + b)^{m-1}} + \frac{2m-3}{2(m-1)b} \int \frac{du}{(u^2 + b)^{m-1}}, \text{ if } m \neq 1.$$

In applying any of the formulas, the following options are possible.

$$\log(u \pm \sqrt{u^2 + a^2}) \text{ or } \sinh^{-1}(u/a).$$

$$\log(u \pm \sqrt{u^2 - a^2}) \text{ or } \cosh^{-1}(u/a).$$

$$\frac{1}{2a} \log \frac{u+a}{u-a} \text{ or } \frac{1}{a} \operatorname{ctnh}^{-1} \frac{u}{a} \text{ if } |u| > |a|.$$

$$\frac{1}{2a} \log \frac{a+u}{a-u} \text{ or } \frac{1}{a} \tanh^{-1} \frac{u}{a} \text{ if } |u| < |a|.$$

$$5. \int u^m \sqrt{au + b} du = \frac{2}{a(2m+3)} [u^{m+1}(au + b)^{\frac{3}{2}} - mb \int u^{m-1} \sqrt{au + b} du].$$

$$6. \int \frac{\sqrt{au + b}}{u} du = 2\sqrt{au + b} \div \sqrt{b} \log \frac{\sqrt{au + b} - \sqrt{b}}{\sqrt{au + b} + \sqrt{b}}, \text{ if } b > 0.$$

$$7. \int \frac{\sqrt{au-c}}{u} du = 2\sqrt{au-c} - 2\sqrt{c} \tan^{-1} \sqrt{\frac{au-c}{c}}, \text{ if } c > 0.$$

$$8. \int \frac{\sqrt{au+b}}{u^m} du = -\frac{1}{(m-1)b} \left[\frac{(au+b)^{\frac{3}{2}}}{u^{m-1}} + \frac{(2m-5)a}{2} \int \frac{\sqrt{au+b}}{u^{m-1}} du \right], \text{ if } m \neq 1.$$

$$9. \int \frac{udu}{\sqrt{au+b}} = \frac{2(au-2b)}{3a^2} \sqrt{au+b}.$$

$$10. \int \frac{u^m du}{\sqrt{au+b}} = \frac{2}{a(2m+1)} \left[u^m \sqrt{au+b} - mb \int \frac{u^{m-1}}{\sqrt{au+b}} du \right].$$

$$11. \int \frac{du}{u\sqrt{au+b}} = \frac{1}{\sqrt{b}} \log \frac{\sqrt{au+b} - \sqrt{b}}{\sqrt{au+b} + \sqrt{b}}, \text{ if } b > 0.$$

$$12. \int \frac{du}{u\sqrt{au-c}} = \frac{2}{\sqrt{c}} \tan^{-1} \sqrt{\frac{au-c}{c}}, \text{ if } c > 0.$$

$$13. \int \frac{du}{u^m \sqrt{au+b}} = -\frac{\sqrt{au+b}}{(m-1)bu^{m-1}} - \frac{(2m-3)a}{(2m-2)b} \int \frac{du}{u^{m-1} \sqrt{au+b}}, \text{ if } m \neq 1.$$

For integrands like $u^n(\log x)^m$, $u^n \sin au$, $u^n \cos au$, $u^n e^{au}$, $u^n \sin^{-1} u$, and $u^n \tan^{-1} u$, with n and m integers > 0 , use integration by parts. Products like $e^{au} \sin bu$ and $\sin bu \sinh cu$ can be handled by twice integrating by parts (see Example 5, Art. 7, Chapter 13).

$$14. \int e^{au} \sin bu \, du = \frac{e^{au}(a \sin bu - b \cos bu)}{a^2 + b^2}.$$

$$15. \int e^{au} \cos bu \, du = \frac{e^{au}(a \cos bu + b \sin bu)}{a^2 + b^2}.$$

$$16. \int \sin u \sinh u \, du = \frac{1}{2}(\sin u \cosh u - \cos u \sinh u) \\ \text{etc.}$$

Integrals of the forms $\sec^{2n+1} u$ and $\csc^{2n+1} u$ are handled in the same way.

$$17. \int \sec^3 u \, du = \frac{1}{2}[\sec u \tan u + \log(\sec u + \tan u)].$$

$$18. \int \csc^3 u \, du = \frac{1}{2}[-\csc u \cot u - \log(\csc u - \cot u)].$$

$$19. \int \sec^{2n+1} u \, du = \frac{\tan u \sec^{2n-1} u}{2n} + \frac{2n-1}{2n} \int \sec^{2n-1} u \, du.$$

$$20. \int \csc^{2n+1} u \, du = -\frac{\cot u \csc^{2n-1} u}{2n} + \frac{2n-1}{2n} \int \csc^{2n-1} u \, du.$$

Integrands of the form $\sec^m u \tan^n u$ and $\csc^m u \cot^n u$, with m odd and n even, can be brought into one of the forms 17-20.

Even powers of $\sin u$ and $\cos u$ can be reduced to multiple angles by repeated use of the formulas $\cos^2 u = \frac{1}{2}(1 + \cos 2u)$ and $\sin^2 u = \frac{1}{2}(1 - \cos 2u)$.

Integrands that are rational functions of $\sin u$ and $\cos u$ can be transformed into rational algebraic fractions by the substitution of $v = \tan(u/2)$,

$$\sin u = \frac{2v}{1+v^2}, \quad \cos u = \frac{1-v^2}{1+v^2} \quad (\text{see pp. 308-310}).$$

Definite integrals that cannot be integrated by formula can be approximated by trapezoidal rule or Simpson's rule, or by the integration of the Taylor approximation.

There are several tables of integrals available, containing hundreds of formulas. Some knowledge of integration is necessary in order to use them intelligently. It is especially important to select the formula which fits the case in hand precisely, and then to identify the given constants and variables with the letters in the formula accurately.

REFERENCES

- A Short Table of Integrals*, Pierce. Ginn and Co., Boston, Mass.
Handbook of Mathematical Tables and Formulas, Burlington. Handbook Publishers, Sandusky, Ohio.
Mathematical Tables from Handbook of Chemistry and Physics, Hodgeman. Chemical Rubber Publishing Co., Cleveland, Ohio.
Tables of Integrals and Other Mathematical Data, Dwight. The Macmillan Co., New York, N.Y.

ANSWERS TO ODD EXERCISES

Chapter 1

Art. 6, p. 11

1. 2; $\frac{3}{2}$.

Art. 8, p. 16

3. $m = -12/x^2$.

7. -2.

9. 1.

11. $m = -1/2(\sqrt{x})^3$.

13. 8π .

Chapter 2

Art. 6, p. 29

4. (a) 2. (c) 1. (e) 1. (g) 0.

5. $k > h, \infty$. $k < h, 0$. $k = h, c_0/b_0$.

Art. 9, p. 36

1. (a) 10. (b) 15. (c) 12. (d) $-\frac{2}{3}$.

5. No limit; no limit.

9. 2, 3.

Art. 12, p. 40

1. (a) all values. (c) $x \neq 1$. (e) all values. (g) $x \neq 1$. (i) all values. (k) $x \neq 0$.

2. (a) 20. (c) 0. (e) 5. (g) -10.

Art. 13, p. 43

1. 1, 0, $-\frac{1}{2}$.

3. 4, -4.

5. -3, -5.

7. -2, $-\frac{1}{4}$.

9. $-\frac{1}{2}$, $-\frac{1}{8}$.

11. -1.

Chapter 3

Art. 5, p. 50

1. $6x^5$.

3. $15x^{14}$.

5. $8x$.

7. $-15x^2$.

9. 1.

11. $16x$.

13. $x^2 + x$.

15. $2ax + b$.

17. $2x$.

19. $-4 - 12x + 15x^2$.

21. 0; -11; 29.

23. 4; 13; 469.

25. 0; -8; 76.

27. $3x - 6y + 2 = 0$.

29. $y + 7 = 0$.

31. 32; 64; 96; 320.

33. 48; 16; -48.

37. .4 tons per mile per hr.

39. (a) $-1/x^2$. (c) $1/2\sqrt{x}$.

40. (a) $-1/(x-1)^2$. (c) $(1+s^2)/(1-s^2)^2$.

Art. 6, p. 54

1. $2/(x+1)^2$, 3. $(x^2+2x-1)/(x+1)^2$, 5. $1/(x-1)^2$,
 7. $2x+2/x^3$, 9. $-1/x^2+2x+3$.

Art. 7, p. 55

1. 32 sq in. per sec. 3. 8π sq in. per sec.
 5. 2 in. per sec; 4 in. per sec. 7. .177 ft per min.
 9. $36\sqrt{3}$ in. per hr.

Art. 8, p. 58

1. $2u(18x^2-6x+4)$, 3. $-6u^2/x^2$,
 5. $-\frac{18}{u^2}(3x-1)(3x^2-2x+1)^2$, 7. $(x+12)(2-x)(3+x)^4$,
 9. $6(x-1)^2/(x+1)^4$.

Art. 11, p. 62

1. (a) $1/2y$, (c) $1/(15y^2+4y)$, (e) $-y^4/12$,
 2. (a) 1, (c) $-\frac{1}{3}$, (e) 3, 3. 90° ; $36^\circ 52'$,
 5. (a) $x/\sqrt{x^2-1}$, (c) $-x/(x^2-1)^{3/2}$, (e) $-x/\sqrt{4-x^2}$.

Art. 12, p. 63

1. $2/(3u)$, 3. $(2t+1)/(3t^2-2t)$, 5. $2u/(1-u^2)$,
 7. $4x-3y=15$, 9. $\alpha=11^\circ 45'$; $\alpha=0$; $\alpha=-30^\circ$.

Art. 13, p. 65

1. (a) $-x/y$, (c) $3x^2/(2y-3)$, (e) $-\sqrt{(y+4)/(x-1)}$,
 2. (a) $-\frac{4}{3}$, (c) $-\frac{8}{3}$, (e) -1 ,
 3. $-\frac{8}{3}$, 5. 10.8 mph.

Art. 14, p. 66

1. $-18x(1-3x^2)^2$, 3. $\frac{3}{2}x^{-\frac{1}{2}}+\frac{1}{2}x^{-\frac{3}{2}}$, 5. $\frac{8}{9}(1-2s)^{-\frac{2}{3}}$,
 7. $\frac{2y}{(1-y^2)^2}\sqrt{\frac{1-y^2}{1+y^2}}$, 9. $\frac{(3-4x^2)^2}{(2x+1)^{\frac{1}{2}}}(3-24x-52x^2)$,
 11. $1/(12y^2)$, 13. $-y^{\frac{1}{2}}/x^{\frac{1}{2}}$, 15. $1/(4y^2+1)$,
 17. $54x/5u^{\frac{1}{2}}$, 19. $12u^{\frac{1}{2}}$,
 21. (a) $2x-y=11$, (c) $5x+y=4$, (e) $x+y=4\sqrt{2}$,
 23. $v=-1$, $a=2$, $t=0$ or $3\frac{1}{2}$; $t=2\frac{3}{8}$, 25. 72 mph.

Chapter 4

Art. 1, p. 70

1. $x^2 - \frac{2}{3}x + 1$. 3. $2/(1-x)^2$. 5. $(2-6x^2)/(1+x^2)^2$.
 7. $-(x+4)/4(x+1)^{3/2}$. 9. $f''' = -24/x^5$. 11. $f''' = -12/(x-1)^4$.

Art. 4, p. 79

1. $v < 0$ for $0 < t < 3$, $s = -1$ when $v = 0$, $a > 0$ for all t .
 3. $v < 0$ for $0 < t < 2$, $s = -15$ when $v = 0$, $a > 0$ for all t .
 5. Starts at $s = 5$ with $v = 0$, $a > 0$. $v > 0$ and $v \rightarrow 1$ with increasing t .

Art. 5, p. 81

1. $y'' = 6$. 3. $y'' = 24t^2$. 5. $y'' = 1/(1+t)^2$.
 7. $y'' = 2t^3/(t^2-1)^2$.

Art. 6, p. 82

1. $y'' = -1/y^3$. 3. $y'' = -81/(25y^3)$. 5. $y'' = 2y^2 - \frac{8}{3}y^3$.

Chapter 5

Art. 3, p. 90

1. (a) (0, 2). (c) (-5, 27). 2. (a) (0, -2). (c) (-\frac{5}{2}, -\frac{1}{2}).
 3. (a) (0, 3) infl. (c) (-3, 19) max, (1, -13) min.
 (e) (-2, 16) max, (0, 0) min, (2, 16) max.
 7. (6, 12) min, (0, 0) max. 9. (\frac{1}{2}, \frac{25}{2}) min, (-4, \frac{1}{2}) max.

Art. 4, p. 96

1. 16×16 . 3. 10, 10. 5. 5, 10.
 7. $\frac{1}{2}$. 9. $6\sqrt{2} \times 6\sqrt{2}$. 11. $3 \times 6 \times 4$
 13. $14 \times 14 \times 28$; 5488 cu in. 15. $4 \times 4 \times 12$.
 17. $12 \times 12 \times 18$. 19. $r = \sqrt{6}$, $h = 2\sqrt{3}$.
 21. width = $2 \times$ height. 23. 7 miles from C.
 27. 3500, \$24.50. 29. $x = (-a + \sqrt{a^2 + 3\alpha\beta - 3b\beta})/3\beta$.

Art. 5, p. 100

1. 30×30 . 3. $r = h = \frac{1}{3}g$. 5. $h = \sqrt{2}r$.
 9. $y/x = (3 - \sqrt{3})/2$. 11. $x/y = \frac{1}{3}g$. 13. $12 \times 12 \times 3$.
 15. $l/w = a/b$. 17. depth/breadth = $\sqrt{3}$. 19. $r = 6.43$, $h = 38.58$.
 21. $h = 4\sqrt{3}$, $w = 2\sqrt{3}$. 23. $r = 6\sqrt{2}$, $h = 24$.

Chapter 6

Art. 3, p. 109 (Constant C omitted)

1. $3x^2/2$.
3. $x^8/8$.
5. $x^3/3 - 3x^2/2 - 5x$.
7. $(x^3 - \sqrt{2x^2} \div x)/2$.
9. $5x^9/4 + 2x^7 - 3x^6 - 3x^5 \div 3x^4/2 - 7x^3 + 4x^2 + 40x$.
11. $-1/(2x^3)$.
13. $2x^{3/2}/3 + 2x^{1/2}$.
15. $-2x^{-1} + 3x^{-2}/2 + 4x + 3x^2/2 - 4x^3/3$.
17. $6x^{5/2}/5 - 4x^{3/2}/3 + x$.
19. $6x\sqrt{x} + 2x - 2x^2\sqrt{x}$.
21. $6x^{3/2} - 6x^{1/2} \div 3x + 3x^{1/2} - 3x^{3/2}$.
23. $3x^{5/2}/2 \div 4x^{1/2} - 4x^{3/2}/3$.
25. $3x^{5/2}/2 + 2x^{3/2} + x/2 + x^2/2 + 2x^{5/2}/3 + 3x^{3/2}/4$.

Art. 4, p. 114 (Constant C omitted)

1. $\frac{2}{3}(x^2 + 1)^{3/2}$.
3. $\frac{1}{8}(3x^4 - 2)^{1/2}$.
5. $-\frac{2}{3}(3 - \sqrt{x})^{3/2}$.
7. $\frac{2}{3}(x - 1)^{3/2}$.
9. $-\frac{1}{18}(1 - x^2)^{3/2}$.
11. $2/(3\sqrt{4 - x^3})$.

Art. 5, p. 116

1. (c) $s = 30t + 10t^2 - 5t^3$.
2. (c) $v = \frac{1}{4}t^4 - 2t^3 + 6t^2 - 5t$.
3. (a) $s = 5t^2 + 4t + 10$. (c) $s = \frac{1}{8}t^3 - 2t^2 + 4t + 10$.
5. $t = 8$, speed = 11.6.

Art. 8, p. 125

1. (a) $v_x = 40$; $v_y = 30 - 32t$. (b) $X = 40t$; $Y = 30t - 16t^2$.
(c) $100Y = 75X - X^2$. (d) $\frac{1}{8}$ sec; 75 ft from 0; $v_x = 40$, $v_y = -30$.
(e) 40 ft/sec.
3. $|v| = \sqrt{5}$.
5. (a) $v_x = 2t - 6$; $v_y = 2$; $a_x = 2$; $a_y = 0$.
(b) $X = 7$; $Y = 14$; $|v| = 2\sqrt{17}$; $= 14^\circ 2'$.
(c) $|a| = 2$; $\beta = 0$.
7. (a) $X = 880t/3$, $Y = -16t^2$. (b) $Y = -9X^2/(220)^2$. (c) 25 sec, 7333 ft forward.

Art. 9, p. 127

1. $y = x/2$.
3. $y = x^2 - x + 3$.
5. $y = 2x^3/3 + 13/3$.
7. $y = 2\sqrt{x} + 1$.
9. $y = (x - 1)^3/3 + 10/3$.

Chapter 7

Art. 3, p. 133

1. 30.
3. 35.
5. 137/60.
7. $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}$.
9. $f(1) + 2f(2) + 3f(3) + 4f(4) + 5f(5)$.
11. $0^2 + 1^2 + 2^2 + 3^2 + \dots + (n-1)^2$.

Art. 4, p. 134

3. $n^2 - n$.

5. $115/462; [1/2 - 1/(n+1)(n+2)]/2$.

Art. 5, p. 136

1. $n^5/5 + n^4/2 + n^3/3 - n/30$.

3. $44i$.

7. 306 .

Art. 6, p. 138

1. $\frac{1}{8}$.

3. $\frac{8}{3}$.

5. 1 .

Art. 8, p. 143

1. $\frac{1}{8}$.

3. $1/a - 1/b$.

5. $\frac{1}{6}$.

7. $\frac{11}{8}$.

Art. 9, p. 148

1. 40 .

3. $\frac{5}{2}$.

5. $\frac{14}{3}$.

7. 8 .

9. 0 .

11. 4 .

13. 9 .

15. $\frac{16}{9}$.

Art. 10, p. 153

1. $-\frac{3}{2}$.

3. -8 .

5. $\frac{59}{3}$.

7. -54 .

9. $12\sqrt{2}$.

11. $\frac{7}{6}$.

Chapter 8

Art. 1, p. 156

1. 2 .

3. $1/\sqrt{3}$.

5. $1 + \sqrt{7}$.

Art. 2, p. 158

1. $(u^2 - 1)/2$.

3. $u^3/3$.

5. $u^3 + u^2/2$.

Art. 4, p. 162

1. 50 .

3. -128 .

5. 70 .

7. 9 .

9. $\sqrt{5} - 1$.

13. $\frac{32}{3}$.

Art. 5, p. 163

1. $x^5 - 2x^4 + 2x^3 + x^2 - 4x + C$.

3. $2\sqrt{x} + 2x^{3/2} + C$.

5. $\frac{1}{2}(ax^3 - bx + c)^2 + C$.

7. $2\sqrt{x^2 + 1} + C$.

Chapter 9

Art. 2, p. 170

1. 36.	3. $\frac{64}{3}$.	5. $\frac{1}{4}$.
7. $\frac{5}{3}$.	9. $\frac{1}{3}$.	11. $2\frac{1}{2}$.
13. 3.	15. $3\frac{1}{2}$.	17. $57\frac{1}{2}$.
19. $a^2/6$.	21. $33\frac{3}{4}$.	23. $\frac{1}{16}$.
25. $8p^2/3$.		

Art. 4, p. 177

1. $\pi/5$.	3. $9\pi/2$.	5. 8π .
7. $111\pi/4$.	9. $(4480\frac{1}{3}\frac{6}{5})\pi$.	11. $(22\frac{2}{3})\pi$.
13. 8π .	15. $3\pi/2$.	
17. (a) $\pi h^2(3a - h)/3$.	(b) $\pi h(3a^2 - h^2)/3$.	19. $4\pi a^2 b/3$.

Art. 5, p. 180

1. 2.	3. $\frac{64}{3}$.	5. 16π .
9. $2hr^2/3$.		

Art. 6, p. 185

1. $25\frac{1}{2}$ ft.-lb.	3. $67\frac{1}{2}$ in.-lb.
5. $40007/41$ mile-tons.	7. $14,243\frac{5}{8}$ ft.-lb.

Art. 7, p. 190

1. 54,000.	3. (a) $8000/3$.	(b) $2389\frac{1}{2}$.
5. 29,250.	7. (a) $468\frac{3}{4}$.	(b) $937\frac{1}{2}$.

Chapter 10

Art. 2, p. 192

1. (a) 2.	(c) $\frac{2}{3}$.	3. 120 in. per sec.
5. (a) .000073 rad per sec.	(b) .29 mi per sec.	(c) .145 mi per sec.

Art. 8, p. 205

21. $-\sin x$.	23. $-a^2 \sin ax$.	25. $18 \cos 6x$.
27. $-a \sin x - b \cos x$.		29. $\cos(x+y)/[1 - \cos(x+y)]$.
31. $\frac{\sin(x-y) - \cos(x+y)}{\sin(x-y) + \cos(x+y)}$.		33. $-\tan t$.
35. $\frac{\sin t - t \cos t}{\cos t + t \sin t}$.	37. 1; 2.	39. 0; 1.

Art. 11, p. 216

1. $x = 2\pi/3$, max., $x = 4\pi/3$, min., 90° . 3. $x = \pi/4$, min., $x = -\pi/4$, max.
 5. 200π ft per sec. $5000\pi/9$ ft per sec. 7. $\sqrt{3}/10$ rad per sec.
 9. 6.67 ft per sec, 5.16 ft per sec. 11. $\pi/12$, $\sqrt{3}\pi/24$, $\sqrt{2}\pi/24$, 0, $-\pi/12$.
 13. 2.72 sq ft per sec. 17. $r = 3\sqrt{2}/2$, $h = 3\sqrt{2}$.

Art. 13, p. 224

7. $x = .5 \sin 200 t$.
 9. (a) $y = c_1 \sin x + c_2 \cos x$. (b) $s = c_1 \sin \sqrt{kt} + c_2 \cos \sqrt{kt}$.

Art. 14, p. 226

1. $\frac{1}{3}$. 3. $\pi/8$. 5. $\frac{4}{3}$.
 7. $\pi/12$. 9. $\pi/2 - 1$. 11. $\pi^2/2$.

Chapter 11

Art. 6, p. 236

21. .43; .86; 4.3; .043. 23. (1, 0); (e, 1).

Art. 9, p. 242

25. 1, $2e$, 0, $-1/e^2$. 27. (0, 1) max; $(\pm \sqrt{5}, e^{-.5})$ inf.

Art. 10, p. 244

1. $y[2x/(x^2 + 1) - 2x/(x^2 - 1)]$.
 3. $y[2x/(x^2 + 1) + x/(x^2 - 1) - 2/(2x + 1)]$.
 5. $y[2 \cot x + 3/(2x + 2)]$. 7. $x^x(1 + \log x)$.
 9. $x^{\sin x}[(\sin x)/x + \cos x \log x]$.

Art. 14, p. 252

3. 0. 5. 0. 11. 0.

Art. 15, p. 254

1. $y = 3e^{2x}$. 3. $y = e^{-4x}$. 5. $y = 2.83e^{-.1x}$.
 7. $y = 2e^{-x}$. 9. $i = 20e^{-.50t}$. 11. $Q = Q_0 e^{.043t}$; $1.19Q_0$.

Art. 16, p. 256

1. 2.30. 3. .999,955. 5. $(e^2 - 3)/2$.
 7. .693. 9. $\frac{1}{2}$. 11. 1.264.
 13. .347. 15. .193. 17. 2.30.
 19. .43. 23. $e - 1$. 25. $\pi(1 - e^{-4})/2$.

Chapter 12

Art. 3, p. 263

1. (a) .96. (c) -.03. (e) .008,686. 2. (a) 9.95. (c) 4.042. (e) .4025.
 3. (a) 70. (c) 2.996. (e) .85725. 5. 45.23.
 7. 62.83; 25.13.

Art. 6, p. 269

1. $(12x^2 - 10x + 2) dx$. 3. $16x(2x^2 - 1)^3 dx$.
 5. $(3 + 2x - 3x^2)/(x^2 + 1)^2 dx$. 7. $[2x/(x^2 - 1) + \cos x] dx$.
 9. $2udu; dx$. 13. $-(t^{-\frac{1}{2}} + 4t^{\frac{1}{2}})$.
 15. $2u/(1 - u^2)$. 17. $(6x - 4xy)/(2x^2 + 3y - 1)$.
 19. $(3y\sqrt{x^2 + y^2} - x)/(y - 3x\sqrt{x^2 + y^2})$.

Art. 9, p. 274

1. 2. 3. $e - 1$. 5. .458.

Art. 10, p. 277

5. (b) 3.912,323.

Chapter 13

(Integration Constants Omitted)

Miscellaneous Exercises, p. 317

1. $-2 \cos \sqrt{x}$. 3. $\log(e^x + e^{-x})$. 5. $\frac{1}{4} \log |(x+1)/(x+5)|$.
 7. $(x+1) \tan^{-1} \sqrt{x} - \sqrt{x}$. 9. $\frac{1}{8}\theta - \frac{3}{4} \sin 2\theta + \frac{1}{2} \sin 4\theta$.
 11. $\frac{1}{2} \sin^{-1}(t^2/3)$. 13. $e^{2t}(2 \sin \pi t - \pi \cos \pi t)/(4 + \pi^2)$.
 15. $(x^2/3) \sin^{-1} x + \frac{1}{3} \sqrt{1-x^2} - \frac{1}{9}(1-x^2)^{\frac{3}{2}}$. 17. $\log |(x-1)^2/x| - x/(x-1)^2$.
 19. $3 \log |\sqrt[3]{x}/(1 + \sqrt[3]{x})|$. 21. $\tan^{-1} e^t$.
 23. $\frac{1}{8}x^4 - \frac{1}{4}x^3 \sin 2x - \frac{3x^2}{8} \cos 2x + \frac{3x}{8} \sin 2x + \frac{3}{16} \cos 2x$.
 25. $e^t/2 - e^t(\cos 2t + 2 \sin 2t)/10$. 27. $\sin^{-1}[(x-2)/3]$.
 29. $\sqrt{9x^2 + 4}(9x^2 + 16)/3 + 8 \log [(\sqrt{9x^2 + 4} - 2)/3x]$.
 31. $17a^2/6$. 33. $3(\tan^{-1} 2)/4$. 35. $\pi/8$.
 37. $\frac{1}{6} + \log 32$. 39. $59 + 6 \log 2$.

- Art. 9, p. 342**

Art. 10, p. 345

1. $a^2/6$. 3. $1 - 1/e^a$. 5. $2\pi a^2$.
 7. $3\pi/2 - 4$. 9. $3\pi a^2/8$. 11. $1 - \log(3/2)$.

Art. 11, p. 350

1. $\pi/2$. 3. $544\sqrt{2}\pi/15$. 5. $2\pi^2$.
 7. 36π . 9. $16\pi(135 - 49\sqrt{5})/15$.
 11. $2a^2\pi^2$. 13. $3\pi/10$. 15. $\pi(1 - 1/e)$.

Art. 12, p. 354

3. $\pi(17\sqrt{17} - 1)/6$. 5. $\pi[12\sqrt{145} + \log(12 + \sqrt{145})]/6$.
 7. $\pi[\sqrt{2} - 5/24 + \log \frac{1}{3}(2\sqrt{3} + \sqrt{6})]$. 9. $64\pi a^2/3$.

Chapter 15

Art. 3, p. 365

1. $\tan \psi = -\cot \theta$; $\tan \alpha = (\tan \theta - \cot \theta)/2$.
 3. 1, ∞ , -1 ; 0. 5. $\pi/2$, $\tan^{-1}2$, $\pi/2$, $\pi/3$. 9. -4.6 .

Art. 4, p. 367

3. $a[\log\{2(\sqrt{\pi^2 + 1} + \pi)/(\sqrt{5} + 1)\} - \sqrt{\pi^2 + 1}/\pi + \sqrt{5}]$.

Art. 5, p. 369

3. $3[\csc(\theta_1/2)]/4a$. 5. $\frac{1}{8}a$.

Art. 6, p. 371

1. 16. 3. π . 5. 18π .
 7. $7\pi^2/3$; $2\pi^3$. 9. $4\pi/3 + 2\sqrt{3}$. 11. $a^2[\sqrt{15} - \cos^{-1} 1/4]/2$.
 15. $a^2(10\pi + 27\sqrt{3})/64$.

Chapter 16

Art. 1, p. 378

1. (a) $1/2$, $\sqrt{3}$, $\sqrt{13}/2$; 0, 2, 2.
 (c) 0, 4, 4; 4, 8, $4\sqrt{5}$; 0, 12, 12.
 2. (a) 0, 10, 1; -5 , $5\sqrt{3}$, 1; -10 , 0, meaningless.
 (c) $-10/\sqrt{2}$, $10/\sqrt{2}$, $1/\sqrt{2}$; $-10/\sqrt{1 + \pi^2}$, $10\pi/\sqrt{1 + \pi^2}$, $\pi^2/\sqrt{1 + \pi^2}$.
 (e) 0, 10, 5; $2\sqrt{5}$, $4\sqrt{5}$, $\sqrt{5}$; 0, 10, $5/3$.

Art. 2, p. 385

1. (a) $0, 2, 4/\sqrt{5}, 2/\sqrt{5}$.
 (c) $-5, 3\sqrt{3}, 8\sqrt{7}/7, 10\sqrt{21}/7; 0, -6, 0, 6$.
 (e) $1, 1, 0, \sqrt{2}; e, 1/e, (e^2 - e^{-2})/\sqrt{e^2 + e^{-2}}, 2/\sqrt{e^2 + e^{-2}}$.
 3. $968 \cos 11t, 968 \sin 11t, (2\pi, 4)$.
 5. $32/5, 8\sqrt{41}/5, 1152/125\sqrt{41}, -288/25\sqrt{41}$.

Art. 3, p. 386

1. $0, 20$.
 3. $2\sqrt{2}, 2\sqrt{2}, 6\sqrt{2}$.

Art. 4, p. 387

1. $54,000$ ft-lb. 3. $18,750\pi$ ft-lb. 5. $63,750\pi$ ft-lb.

Art. 5, p. 393

1. $2kM/a\sqrt{L^2 + 4a^2}$. 3. $kM[1/\sqrt{b^2 + a^2} - 1/\sqrt{(b+L)^2 + a^2}]/L$.
 5. $\pi k\delta(2 - \sqrt{2})$.

Art. 6, p. 399

5. $2r^3/3$. 7. $8p^3/5$.
 9. (a) $a^2b/6$; (b) $a^2b/3$. 11. $\pi r^4/4$.

Art. 7, p. 404

1. $4r/3\pi$ from the center. 3. $(3p/5, 3p/4)$. 5. $h/4$ from the base.
 7. $(4a/3\pi, 4h/3\pi)$. 9. $(\pi a, 5a/6)$. 11. $\bar{x} = 3a/8$.
 13. $(\frac{3}{4}, \frac{1}{4})$. 15. $\bar{y} = (e^2 \div 4 - e^{-2})/4(e - e^{-1})$.

Art. 8, p. 407

1. $\pi a^2/3$. 3. $bh^2/12, bh^3/4$.
 5. $64/105$. 7. $5000, 1666\frac{2}{3}$.

Art. 9, p. 409

1. $\pi hr^4/10$. 3. $8\pi/3$. 5. $\pi(1 - e^{-4})/8$.

Chapter 17

Art. 2, p. 414

1. 3 . 3. $e - \frac{1}{2}$. 5. $\log \sqrt{3}$.

Art. 3, p. 419

1. $\pi/2$.	3. div.	5. div.
7. div.	9. $1/(1-a)$ if $a < 1$, div. if $a \geq 1$.	

Art. 4, p. 421

1. $\frac{1}{2}$.	3. div.	5. div.
7. $\pi/12$.	9. $\sqrt{2}/2$.	11. div.
13. div.		

Art. 5, p. 424

1. conv.	3. conv.	5. conv.
7. div.	9. div.	11. conv.
13. div.	15. conv.	17. div.
19. conv.	23. $-\pi$.	

Chapter 18

Miscellaneous Exercises, p. 435

1. $4 \sinh 4x$.	3. $\sinh 2x$.
5. $\sin x \sinh x + \cos x \cosh x$.	7. $\cosh x$.
9. $\operatorname{sech} x$.	13. $\frac{1}{2} \sin^{-1}(2x/3) + C$.
15. $\frac{1}{5} \cosh^5 x - \frac{1}{3} \cosh^3 x + C$.	17. $\frac{1}{4} e^{2x} - \frac{1}{2} x + C$.
19. $-(1/x) \sinh^{-1} x - \tanh^{-1} \sqrt{1-x^2} + C$.	

Chapter 19

Art. 2, p. 438

1. 21,250.	3. 1.19.	5. 10.06.

Art. 3, p. 442

1. 21,222.	3. 1.176.	5. 6.66667; 6.63017.
7. 7.938; 31.752.	9. 37.33.	11. 4367.4 cu yd.
13. 12,666 $\frac{2}{3}$ cu ft.	15. 3828.8 ft-lb.	

Chapter 20

Art. 4, p. 452

1. .1823; $R < .000045$.	3. 9; 99.	7. 2.

Art. 6, p. 460

13. $-x^2/2 - x^4/12 - (x^6/90) [15 \sec^6(\theta x) - 15 \sec^4(\theta x) - 2 \sec^2(\theta x)]$.

Art. 7, p. 463

1. .479427.

3. .322.

5. 1.5431.

Art. 8, p. 467

1. 0.

3. 1.

5. $-\frac{1}{8}$.7. $\frac{1}{2}$.

9. 0.

11. 0.

13. 1.

15. 0.

17. -1 .19. $\sqrt{e}$.21. $\frac{3}{2}$.

23. 5.

25. $\frac{1}{2}$.

27. 2.

29. $1/e$.

31. 0.

33. $-\pi$.35. $\frac{1}{2}$.37. $\frac{3}{2}$.

39. 1.

41. $\pm \infty$.43. e .

45. 0.

Chapter 21

Art. 3, p. 474

2. (a) $1/2i$. (c) $(i-1)/i$. (e) $(-2)^{i-1}/i!$.

Art. 5, p. 480

3. Conv.

5. Div.

7. Conv.

9. Div.

11. Div.

13. Div.

15. Conv.

16. (a) Div. (c) Conv.

Art. 6, p. 483

1. Div.

3. Conv.

5. Div.

9. $\sum i^n/e^i$ conv.

10. (a) 2; (c) 11.

Art. 7, p. 486

1. Conv.

3. Conv.

5. Conv.

7. Div.

9. Div.

11. Div.

13. Conv.

15. Conv.

Art. 9, p. 489

1. Cond.

3. Cond.

5. Cond.

7. Abs.

9. Cond.

11. (1) $\frac{1}{2}$. (3) $\frac{1}{2}$. (5) $1/\log 5$. (9) $1/(5 \log 5)$.

12. (a) .947. (c) .465.

Miscellaneous Exercises, p. 492

1. Conv.

3. Conv.

5. Div.

7. Conv.

9. Conv.

11. Abs. conv.

13. Conv.

15. Div.

Chapter 22

Art. 2, p. 497

1. $-1 < x < 1$. 3. $-1 \leq x < 1$. 5. $-1 < x < 1$.
 7. $-1 \leq x \leq 1$. 9. $1 < x \leq 3$. 11. $-\frac{1}{3} \leq x \leq \frac{1}{3}$.
 13. $-10 < x < 10$. 15. $0 < x \leq 2$.

Art. 3, p. 499

1. $1 + 2x + 4x^2 + 8x^3 + \dots$, $|x| < \frac{1}{2}$.
 3. $\frac{1}{a} \left(1 - \frac{b}{a}x + \frac{b^2}{a^2}x^2 - \frac{b^3}{a^3}x^3 + \dots \right)$, $|x| < \frac{a}{b}$.
 9. $\log a + \frac{b}{a}x - \frac{b^2}{2a^2}x^2 + \frac{b^3}{3a^3}x^3 - \dots$, $-\frac{a}{b} < x \leq \frac{a}{b}$.

Art. 4, p. 501. (Assume $0! = 1$.)

1. $\sum_0^\infty x^i/i!$, all values. 3. $\sum_0^\infty (-1)^i x^{2i}/(2i)!$, all values.
 5. $\sum_0^\infty x^{2i}/(2i)!$, all values. 7. $\sum_1^\infty (-1)^{i-1} (2x)^i/i$, $-\frac{1}{2} < x \leq \frac{1}{2}$.
 9. $\sum_0^\infty 1 \cdot 3 \cdot \dots \cdot (2i-1)x^i/(i! \cdot 2^i)$, $|x| < 1$.
 11. $x/2 + x^2/8 - x^4/192 + x^5/480$. 13. $\log 2 - x^2/4 - x^4/96 + x^6/1152$.
 15. $x - x^2/2 - x^3/6 - x^4/12$.
 17. $e^2[1 + (x-2) + (x-2)^2/2! + (x-2)^3/3!]$.
 19. $3.76 + 3.62(x-2) + 3.76(x-2)^2/2! + 3.62(x-2)^3/3!$
 21. $\sum_0^\infty (-1)^i x^{2i}/i!$. 23. $\sum_0^\infty (-1)^i x^{2i}/(2i+1)!$.
 25. $u - u^2/(2 \cdot 3) + 1 \cdot 3u^3/(2 \cdot 4 \cdot 5) - 1 \cdot 3 \cdot 5u^5/(2 \cdot 4 \cdot 6 \cdot 7) + \dots$.

Art. 7, p. 508

1. $1 + x - x^2/2! - x^3/3! + x^4/4! + x^5/5! + \dots$, all values.
 3. $1 + x - x^3/3 - x^4/6 - x^5/30 + \dots$, all values.
 5. $x \log 2 + x^2/2 - x^3/8 + x^4/24 - x^5/64 + \dots$, $-2 < x \leq 2$.
 7. $x^2/2 + x^3/3 \cdot 1! + x^4/4 \cdot 2! + x^5/5 \cdot 3! + x^6/6 \cdot 4! + \dots$, all values.
 9. $x - x^3/3 \cdot 1! + x^5/5 \cdot 2! - x^7/7 \cdot 3! + x^9/9 \cdot 4! - \dots$, all values.

Art. 8, p. 511

1. $1 + x + x^2/2! + x^3/3! + \dots (= e^x)$.
 3. $2 + x - x^2/4 + x^3/8 - 5x^4/64 + \dots (= 2\sqrt{1+x})$.
 5. $5 + 4x + 2x^2 + x^3/2 + 5x^4/24 - \dots$.
 7. $2[1 + x^2/2! + x^4/4! + x^6/6! + \dots] (= 2 \cosh x)$.

Chapter 23

Art. 3, p. 516

5. $y^2 \div (z - 3)^2 = 9$. 7. $4y = x^2$. 9. $y = z = 0$.

Art. 4, p. 517

1. (a) $\sqrt{65}$, $\sqrt{58}$, 5; (c) 20, $\sqrt{337}$, 15. (e) $\sqrt{89}$, 10, $\sqrt{61}$.

Art. 5, p. 518

1. (a) $\sqrt{74}$. (c) $\sqrt{481}$. (e) $\sqrt{125}$.
 3. $\sqrt{73}$, $\sqrt{206}$, $\sqrt{182}$, $\sqrt{41}$, $\sqrt{117}$, $\sqrt{164}$. 5. $2x + 2y + 2z = 15$.
 7. $x^2 \div y^2 \div z^2 = 100$. 9. (1, 1, 11), (7, 7, 2), (-3, -4, -10).

Art. 6, p. 519

1. $x^2 \div y^2 \div z^2 \div 4x - 2y - 6z - 2 = 0$.
 3. (2, -3, -2), rad = 8.
 5. $f(x, y) = -2 - \sqrt{64 - (x - 2)^2 - (y + 3)^2}$.

Art. 7, p. 521

3. $z = 1 - (x^2 + y^2)$.
 5. Two paraboloids: $z = \pm (x^2 + y^2)$, axis on z -axis.
 7. Paraboloid with axis on x -axis.
 9. Cone of revolution about z -axis.
 11. $x^2 \div z^2 = 3y^2$.

Art. 8, p. 524

1. $x = ay^2 \div bz^2$; $y = ax^2 \div bz^2$.

Art. 9, p. 527

1. (a) $1/\sqrt{3}$, $1/\sqrt{3}$, $1/\sqrt{3}$; (c) $2/\sqrt{13}$, $3/\sqrt{13}$, 0. (e) 0, 0, 1.
 7. 1, 0, 0; 0, 1, 0; 0, 0, 1. 9. (a), (c), (d), (e).

Art. 10, p. 529

5. (a) and (c).

Art. 11, p. 532

1. $5/\sqrt{30}$, $-1/\sqrt{30}$, $2/\sqrt{30}$. 3. $x + y + z = 6$.

Art. 12, p. 537

1. $2x = -2y = \sqrt{2}z$. 3. $\frac{x-5}{3} = \frac{y-2}{5} = \frac{z+1}{-5}$.
 7. $\frac{x-1}{-1} = \frac{y-1}{2} = \frac{z-1}{3}$. 9. $\frac{x+3}{5} = \frac{y-4}{-8} = \frac{z+5}{1}$.

Art. 13, p. 541

1. Parabola.

3. Parabola.

5. Oval-shaped curves, one above the xy -plane, one below, cut from the sphere by the cylinder.

7. $\frac{x-1}{2} = \frac{y-1}{3} = \frac{z-1}{4}$; $2(x-1) + 3(y-1) + 4(z-1) = 0$.

9. $2\pi\sqrt{a^2 + b^2}$.

11. $\frac{x-3}{1} = \frac{y-4}{-\frac{3}{4}} = \frac{z-9}{6}$.

Art. 14, p. 543

5. (a) $r = 10$; (b) $r = 2 \cos \theta$.7. (a) $r = a$; (b) $r = 2a \cos \theta$.

9. $r^2 = 3z^2$.

Chapter 24

Art. 2, p. 551

1. $0, 1; \frac{4}{3}, -\frac{3}{5}; -\frac{1}{13}, \frac{1}{13}; 1, 0$.

3. $9, 1; -9, -1; 0, 0; 45, 29$.

5. $0, 0; 0, 0; -\frac{1}{18}, \frac{1}{18}; -\frac{1}{18}, \frac{1}{18}$.

7. $\frac{1}{3}, \frac{1}{2}; \frac{2}{3}, \frac{4}{3}; \frac{1}{3}, \frac{2}{3}; \frac{1}{3}, \frac{2}{3}; \frac{2}{3}, \frac{1}{3}$.

9. $0, \frac{1}{2}; -\frac{1}{2}, \frac{1}{2}; -\frac{1}{2}, -\frac{1}{2}$.

11. $-\frac{4}{5}$.

15. $-1, 1, 2; -\frac{1}{2}, \frac{1}{2}, 0$.

17. $-3e^2, 6e^2, 4e^2$.

19. $0, 2x/y^3, -1/y^2$.

21. $0, -x \sin y, \cos y$.

Art. 4, p. 557

1. $6x + 8y - 5z = 25$; $\frac{x-3}{6} = \frac{y-4}{8} = \frac{z-5}{-5}$.

3. $8x + 6y + z = 28$; $\frac{x-1}{8} = \frac{y-3}{6} = \frac{z-2}{1}$.

5. $4x + 3y - 10z = 125$; $\frac{x-4}{4} = \frac{y-3}{3} = \frac{z+10}{-10}$.

$4x + 3y + 10z = -125$; $\frac{x+4}{4} = \frac{y+3}{3} = \frac{z+10}{10}$.

7. $9x + 16y + 25z = -25$; $\frac{x+2}{9} = \frac{y+2}{16} = \frac{z-1}{25}$.

Art. 5, p. 560

1. $2 + 3\sqrt{3}$.

3. $(\sqrt{3} - \frac{1}{2})e^{-2}$.

Art. 6, p. 562

1. -25 .

5. 6.

11. $\sqrt{21}$.

Art. 7, p. 565

1. $(6xy + 2y^2) dx + (3x^2 + 4xy) dy$.
3. $dx/x + dy/y + dz/z$.
5. $e^y(dx + x dy)$.
7. $e^{-x}(-\sin y dx + \cos y dy)$.
9. $(x dy - y dx)/(x^2 + y^2)$.
11. $(x dx + y dy)/\sqrt{x^2 + y^2}$.
13. 580π cu. in.
15. 1.9992; .0068.
17. 28.003.

Art. 8, p. 569

1. $2(x \cos t - ye^t)/(x^2 - y^2)$.
3. $2pq/(1 + x^2), (p^2 - \cos q)/(1 + x^2)$.
5. $s/[(t + s)\sqrt{ts}], -t/[(t + s)\sqrt{ts}]$.
7. $\cos \sqrt{x^2 + y^2} \cdot (x + 2yt)/\sqrt{x^2 + y^2}, 2 \cos \sqrt{x^2 + y^2} \cdot (x - ys)/\sqrt{x^2 + y^2}$.
9. $e^0(2x - y), e^0(2y + x)$.

Art. 10, p. 573

1. $-\pi$.
3. 0.
5. $-z/x, -z/y$.
7. $xz/(1 - z^2), z(2y^2 - 1)/[y(2 - 2z^2)]$.
9. $x + 2y + z = 2, 2(x - 1) = y = 2(z - 1)$.
11. $xx_0/a^2 + yy_0/b^2 + zz_0/c^2 = 1, \frac{a^2(x - x_0)}{x_0} = \frac{b^2(y - y_0)}{y_0} = \frac{c^2(z - z_0)}{z_0}$.

Art. 11, p. 576

1. (a) $x - y = 4, z + 1 = 0, x + y = 0$.
- (b) $\frac{x - x_0}{x_0(x_0^2 - y_0^2)} = \frac{y - y_0}{y_0(x_0^2 - z_0^2)} = \frac{z - z_0}{z_0(y_0^2 - x_0^2)}$
 $x_0(z_0^2 - y_0^2)x + y_0(x_0^2 - z_0^2)y + z_0(y_0^2 - x_0^2)z = 0$.
5. $\frac{x - 2}{2} = \frac{y + 2}{1} = \frac{z - 4}{-1}; 2x + y - z + 2 = 0$.
7. $\frac{x - 5}{10} = \frac{y - 5}{-10} = \frac{z - \pi/8}{-1}; 10x - 10y - z + \pi/8 = 0$.

Chapter 25

Art. 3, p. 582

1. $170\frac{2}{3}$.
3. $\frac{4}{3}$.
5. 0.
7. $\pi/2 - \pi/8$.
9. 1.
11. 8.
13. $37\frac{1}{2}$.
15. $2 \log 2 - 1$.
17. $4\sqrt{3} + 8\pi/3$.
19. $(2\sqrt{2} - 2)/3$.

Art. 4, p. 589

1. $10\frac{3}{4}$.	3. $16\sqrt{2}/15$.	5. $4\frac{1}{2}$.
7. 1.	9. $\frac{1}{2}$.	11. $abc/6$.
13. 3π .	15. $\frac{1}{6}$.	

Art. 5, p. 594

1. 0.	3. $\frac{32}{45}$.	5. $a^2(8\pi + 3\sqrt{3})/12$.
7. $\frac{1}{3}$.	11. 5.4 .	

Art. 6, p. 598

1. $8\sqrt{3}/3$.	3. $2\sqrt{2}\pi$.	5. $8a^2$.
7. $\pi[(4a^2 + 1)^{\frac{3}{2}} - 1]/6$.	9. $16a^3$.	

Art. 7, p. 600

1. $2kb^4/3$; $(5b/8, 5b/8)$.	3. $kb^3/6$; $(b/4, 2b/3)$.
5. $4ka$; $(a/3, 0)$.	7. $kb^5\sqrt{2}/120$; $kb^5\sqrt{2}/120$; $kb^5\sqrt{2}/8$.
9. $k \tan^{-1}(1/2)$.	11. $2\pi k[1 - b/\sqrt{a^2 + b^2}]$.

Art. 9, p. 605

1. $17\frac{1}{8}$.	7. $-\pi/8$.	9. $k(e^4 - 1)^3$.

Art. 11, p. 614

3. $\pi k h^4[(a^2 + h^2)^{\frac{3}{2}}/h^3 - 1]/6$.	5. $ka^2b^2c^2/6$.
7. $8a/15, 16a/15\pi, 16a/15\pi$.	9. $\pi ka^5h/15$.
11. $\pi h k[2h^5 + (3a^2 - 2h^2)(h^2 + a^2)^{\frac{3}{2}}]/45$.	13. $\pi ka^3/4$.
15. $\pi ka^2(\sec \alpha - 1)$.	
17. $2k\pi[h + \sqrt{b^2 + a^2} - \sqrt{(b+h)^2 + a^2}]$.	

Index

- Absolute convergence, 487
- Absolute value, 7
- Acceleration, 75
 - in curvilinear motion, 118, 379, 385
 - components in polar coordinates, 385
 - components in rectangular coordinates, 118, 379
- Alternating series, 488
- Amplitude, 195
- Angles, direction, 524
- Angular velocity, 374
- Antiderivative, 105
- Approximation(s), and the mean value theorem, 275
 - by a differential, 261, 275
 - by polynomials, 446
 - by Taylor's theorem, 461
 - of area, 16
 - of functional values, 258
 - of functions, 446
 - of integrals, 436
 - of logarithms, 451
 - of work, 18
 - to the slope, 10
- Arc length, as a function, 337
 - in polar coordinates, 365
 - in space, 540
- Area, in polar coordinates, 369
 - of a surface, 595
 - of a surface of revolution, 350
 - under a curve, 137, 139, 141
 - calculation of, 164
- Area problem, 9, 16, 130, 343
 - in polar coordinates, 369
- Attraction of gravity, 388, 601
- Base of logarithmic function, 232
 - change of, 233
- Bounded set, 25
- Burlington, R. S., 619
- Cardioid, 330
- Catenary, 433
- Cauchy, A. L., 46, 475
- Cauchy test for convergence, 475
- Center of gravity, 400, 599, 611
- Centripetal force, 381
- Centroid, 400
- Chain rule, 57
 - for several variables, 566
- Churchill, R. V., 511
- Circular functions, 426
- Comparison of order, 250
- Comparison test for convergence or divergence, 477
- Comparison with an integral, 481
- Components, of acceleration, 118, 379
 - of velocity, 117, 373
- Compound functions, 56, 566
- Concavity, direction of, 74
- Constant, 2
 - arbitrary, in inverse differentiation, 106
- Continuity, 37
 - of algebraic functions, 39
 - of functions of two variables, 546
 - of power series, 503
 - of $\sin x$, 196
 - of trigonometric functions, 197
- Continuous functions, 37, 546
 - theorems on, 37
- Convergence, of improper integrals, 416
 - of power series, interval of, 495
 - uniform, 502
 - of a sequence, 25
 - theorems on, 25, 475

- Convergence, of a series, 471, 477
 absolute, 487
 conditional, 487
 see also Sequence; Series; Tests
 Coordinates, cylindrical, 541, 606
 polar, 355
 rectangular, 5, 512
 spherical, 542, 608
 Courant, R., 2, 102, 147, 233, 445, 511, 570, 578, 581
 Critical point, 89
 Critical value, 89
 Curvature, 338
 center of, 341
 circle of, 341
 in polar coordinates, 367
 radius of, 341
 Curve, in implicit form, 572
 in parametric form, 327
 in space, 537, 574
 length of, 333, 365, 540
 Curve tracing, 319-332
 Curves, intersection of, 320
 Curvilinear motion, 116, 379
 in polar coordinates, 373
 Cusp, 331
 Cycloid, 331
 Cylindrical coordinates, 541, 606
 Cylindrical surface, 515

 Decreasing function, 87
 Definite integral, 144
 approximation of, 436
 existence of, 147
 Dependent variable, 4
 Derivative, 44
 first, 68
 higher-order, 69
 in a given direction, 557
 mixed, 551
 notation for, 46
 of a power, 48
 of a product, 52
 of a quotient, 53
 of a sum, 49
 of sine function, 202
 ordinary, 548
 partial, 548
 second, 68

 Derived function, 45
 first, 68
 notation for, 46
 second, 68
 Difference quotient, 260
 Differences, 260
 Differentiability, 265, 558
 Differentiable function, 265, 555
 Differential, 261, 563
 approximation by, 261, 275, 564
 exact, 564
 of a function of several variables, 563
 total, 564
 Differential calculus, 45
 Differential coefficient, 267
 Differential equation, 128, 221, 253
 power series solution of, 508
 Differentiation, 45
 of exponential functions, 240
 of hyperbolic functions, 430
 of inverse trigonometric functions, 208
 of logarithmic functions, 234
 of power series, 505
 of trigonometric functions, 202, 203
 see also Derivative
 Direction angles, 524
 Direction cosines, 524, 528
 Direction numbers, 526
 Directional derivative, 558
 Discontinuous function, 37
 Discontinuous trigonometric functions, limits of, 198
 Distance along a curve, 120, 333
 in three dimensions, 516
 Divergence, of improper integrals, 416
 of sequences, 25
 of series, 471, 477
 see also Sequences; Series; Tests
 Domain integral, 578, 601
 Double integral, 578
 Dwight, H. B., 311, 619

 e , 233
 Ellipsoid, 522
 Error(s), in approximation by convergent alternating series, 489
 in differential approximations, 276

- Error(s), in polynomial approximations, 448, 459
 in Simpson's rule, 445
 in trapezoidal rule, 445
 Euler's theorem on homogeneous functions, 553, 570
 Exponential functions, 238
 differentiation of, 240
 inverse differentiation of, 247
 Function, 3
 circular, 426
 compound, 56, 566
 continuous, 37, 546
 decreasing, 87
 defined by an integral, 156, 227
 derived, 45
 differentiable, 265, 555
 discontinuous, 37
 double-valued, 4
 exponential, 238, 240
 expressed by a power series, 494
 homogeneous, 553
 hyperbolic, 428
 implicit, 63, 570
 increasing, 87
 inverse, 60
 inverse hyperbolic, 429
 inverse trigonometric, 206
 limit of, 31
 logarithmic, 230
 monotonic, 59
 of two or more variables, 545
 parametric, 62
 periodic, 194
 potential, 612
 primitive, 105
 single-valued, 4
 trigonometric, 193, 197
 Functional equation, 9, 128
 Functional notation, 5
 Functional relation, 2, 3, 5
 Fundamental theorem of calculus, 158
 Fundamental theorem on primitive functions, 277
 Geometric series, 472
 Gradient, 560
 Gravitational attraction, 388
 Gravity, center of, 400, 599, 611
 Half-life, 254
 Harmonic series, 473, 474, 480
 Higher derivatives, 69
 of implicit functions, 81
 of parametric functions, 79
 Hodgeman, C. D., 619
 Homogeneous function, 553
 Hooke's law, 182, 222
 Hyperbolic functions, 428
 in exponential form, 435
 properties of, 431
 Hyperbolic paraboloid, 522
 Hyperboloid, 523
 Hypocycloid, 332
 Implicit equation, 63
 Implicit function, 63, 570
 Improper integral, 416
 tests for convergence of, 422
 Increasing function, 87
 Indefinite integral, 157
 Independent variable, 4
 Indeterminate form, 464
 Inertia, moment of, 405, 600, 612
 Infinite decimals, 490
 Infinite sequence, 21
 See also Sequence(s)
 Infinite series, 470
 See also Series
 Infinitely many, 21
 Infinity, tending to, 22
 Inflection point, 74, 89, 321
 Instantaneous speed, 12
 Integral(s), 146
 approximation of, 436
 definite, 144
 domain, 578, 601
 double, 578
 existence of, 147
 function defined by, 156, 227
 improper, 416
 indefinite, 157, 163
 iterated, 579, 602
 of function with a finite discontinuity, 411

- Integral(s), of function with an infinite
 discontinuity, 415
 over an infinite range, 419
 theorems on, 149
 triple, 601
 Integrand, 146, 279
 Integration, 146
 approximate, 436
 by partial fractions, 289
 by parts, 295
 by tables, 311
 by trigonometric substitution, 303
 formulas for, 282, 432, 616
 limits of, 146
 of exponential functions, 247
 of hyperbolic functions, 433
 of power series, 503
 of trigonometric functions, 217, 299
 Interval, 7
 closed, 7
 of convergence of power series, 495
 open, 7
 Inverse derivative(s), 105
 of $1/x$, 244
 See also Integration
 Inverse forms of a functional relation,
 59
 Inverse functions, 60
 Inverse logarithmic function, 237
 Inverse operations, 104
 Inverse rate problem, 104
 Inverse trigonometric functions, 206,
 208
 differentiation of, 208
 principal values of, 206, 208
 Involute of the circle, 337
 Iterated integral, 579, 602
 in cylindrical coordinates, 606
 in polar coordinates, 590
 in spherical coordinates, 608
 triple, 603
 Knopp, K., 239, 492, 511
 Lagrange, 46
 Leibniz, 46
 Lemniscate, 371
 Length of a curve, 333, 540
 Level points, 89
 L'Hospital's rule, 464
 Limaçon, 365
 Limit, of a function, 31
 theorems on, 33
 of a function of several variables,
 546
 of a sequence, 23
 theorems on, 25, 475
 Limits, of discontinuous trigonometric
 functions, 198
 of integration, 146
 Liquid pressure, 186
 Logarithmic function, 230
 base of, 232
 differentiation of, 234
 Logarithmic spiral, 362
 Logarithms, approximation of, 451
 laws of, 229
 Napierian, 233
 natural, 233
 Maclaurin's series, 501
 Mass, calculation of, 598
 Mass-point, 17, 388
 Maximum, 83
 of a function of two variables, 560
 relative, 83, 86, 87
 Maximum points, 321
 Mean value theorem, for derivatives,
 270, 271
 for integrals, 155
 Milne, W. E., 445
 Minimum, 83
 of a function of two variables, 566
 relative, 83, 86, 87
 Minimum points, 321
 Mixed derivative, 551
 Moment, of force, 394
 of inertia, 405, 407, 600, 612
 of mass, 599
 static, 398
 Monotonic function, 59
 Monotonic nondecreasing sequence, 476
 Motion, curvilinear, 116
 rectilinear, 74, 114
 simple harmonic, 224
 uniform, 75

- Napier, 233
 Napierian logarithms, 233
 Newton's first law of motion, 116

 Octant, 513
 Order, of tending to infinity, 250
 of vanishing, 201, 250
 Ordinary derivative, 548

 Paraboloid, elliptic, 522
 hyperbolic, 522
 of revolution, 520
 Parameter, 62
 Parametric equations, 62, 79
 of a curve, 327, 537
 of a line, 535
 Partial derivative, 548
 Partial fractions, 287
 integration by, 289
 Period, 194
 Periodic function, 194
 Phase, 224
 Pierce, B. O., 619
 Planes, equations of, 529
 Polar coordinates, 355
 acceleration in, 385
 arc length in, 365
 area in, 369
 curvature in, 367
 curvilinear motion in, 373
 iterated integral in, 590
 tangent problem in, 357
 velocity in, 373
 Polar curve, 356
 Polar moment of inertia, 407
 Pole, 355
 Polynomial approximations, 446
 Potential, 612
 Potential function, 612
 Power formula, for differentiation, 48
 for inverse differentiation, 110
 for rational exponents, 61
 Power series, 494
 continuity of, 503
 differentiation of, 505
 for known functions, 498
 function defined by, 494
 integration of, 503
 Power series, interval of convergence
 of, 495
 limitations of, 511
 theorems on, 503
 uniform convergence of, 502
 uniqueness of, 506
 Power series solutions of differential
 equations, 508
 Pressure, liquid, 186
 Primitive functions, 105, 162
 fundamental theorem on, 277
 Principal values of inverse trigono-
 metric functions, 206, 208
 Product formula for differentiation, 52

 Quadric surfaces, 521
 Quotient formula for differentiation, 53

 Radian, 192
 Radius of curvature, 341
 Radius vector, 357
 Range, of independent variable, 4
 of independent variables, 545
 Rate, 13
 Rate problem, 11
 Rates, related, 54
 Ratio test for convergence, 484
 Rectangular coordinates, 5
 in three dimensions, 512
 Rectifiable curve, 336
 Rectilinear motion, 114
 Reduction formulas, 311
 Related rates, 54
 Remainder, in polynomial approxima-
 tion, 448
 in Taylor's theorem, 455
 Representation of numbers by infinite
 series, 490
 Revolution, figures of, 172
 paraboloid of, 570
 surface of, 350, 520
 volume of, 173, 346
 Robbins, H., 102
 Rolle's theorem, 271
 Root test for convergence, 485

 Sequence(s), 21
 convergent, 25

- Sequence(s), divergent, 25
 limit of, 23
 monotonic, 476
 of sums, 139, 471
 tending to infinity, 22
 theorems on, 25, 475
- Series, alternating, 488
 convergence of, 471
 function expressed by, 494
 geometric, 472
 harmonic, 473, 474, 480
 infinite, 470
 Maclaurin's, 501
 power, 494
 representation of numbers by, 490
 sum of, 471
 summation of, 492
 Taylor's, 458, 500
- Shell method for volumes of revolution, 346
- Simple harmonic motion, 224
- Simpson's rule, 439
 error in, 445
- Slice formula, 173, 178
- Slope of a tangent, 9
 defined, 42
- Space curve, 537, 574
- Speed, 11
 average, 12
 instantaneous, 12
- Sphere, equation of, 518
- Spherical coordinates, 542, 608
- Spiral, logarithmic, 362
 of Archimedes, 365
 reciprocal, 365
- Spring constant, 182, 222
- Static moment, 398
- Stationary points, 89
- Straight lines, equations in three dimensions, 532
- Summation, 132
 of series, 492
- Surface, area of, 595
 in implicit form, 573
 of revolution, 350, 520
- Tangent, 10
 defined, 31, 42
- Tangent, slope of, 9, 42
- Tangent plane, equation of, 554
 existence of, 554
- Tangent problem, 9, 20
 and curve tracing, 319
 in polar coordinates, 357
 in three dimensions, 547
 intuitive solution of, 14
 solution of, 30, 41
- Taylor, Brook, 455
- Taylor's expansion, 458
- Taylor's series, 500
 with remainder, 458
- Taylor's theorem, 455
 approximations by, 461
 remainder in, 455
- Tests, for convergence of improper integrals, 422
 for convergence of series, alternating series, 488
 comparison tests, 477
 integral test, 481
 ratio test, 484
 root test, 485
 for maxima and minima, 321
- Torque, 394
- Trajectory, 122
- Trapezoidal rule, 437
 error in, 445
- Trigonometric functions, 193, 197
 differentiation of, 202, 203
 integration of, 217, 299
- Triple integral, 601
- Uniform convergence of power series, 502
- Uniform motion, 75
- Variable(s), 2
 change of, in differentiation, 57, 566
 in integration, 279
 dependent, 4
 independent, 4
 range of, 4, 545
- Vector, 118
 radius, 357
- Velocity, at a point, 118
 in curvilinear motion, 117, 373

Velocity, in polar coordinates, 373
Volumes, of revolution, 173
 by the shell method, 346
 moment of inertia of, 408
 of special solids, 171, 178
 under a surface, 577

Volumes, under a surface, calculation
 of, 588, 593
Work, by variable force, 182
 in pumping a liquid, 386
Zeno's race-course paradox, 470

Downloaded from www.dbraulibrary.org.in